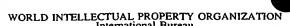
BEST AVAILABLE COPY









INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6:

C12N 1/21, 5/10, 9/10, 15/53, 15/54, 15/61, 15/63, C12P 23/00, C12Q 1/68 (11) International Publication Number:

WO 99/63055

9 December 1999 (09.12.99) (43) International Publication Date:

(21) International Application Number:

PCT/US99/12121

A1

(22) International Filing Date:

2 June 1999 (02.06.99)

(30) Priority Data:

09/088.724

2 June 1998 (02.06.98)

US

09/088,725

2 June 1998 (02.06.98)

US

(71) Applicant (for all designated States except US): UNIVERSITY OF MARYLAND [US/US]; Office of Technology Liaison, 4312 Knox Road, College Park, MD 20742 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): CUNNINGHAM, Francis, X., Jr. [US/US]; 2727 Washington Avenue, Chevy Chase, MD 20815 (US). SUN, Zairen [US/US]; 3405 Tulane Drive #22, Hyattsville, MD 20783 (US).

(74) Agents: GOLDHUSH, Douglas, H. et al.; Nikaido, Marmelstein, Murray & Oram LLP, Suite 330 - G Street Lobby, Metropolitan Square, 655 Fifteenth Street, N.W., Washington, DC 20005-5701 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD. GE. GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report. With amended claims.

(54) Title: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND METHODS OF USE THEREOF

(57) Abstract

Nucleic acid sequences encoding ϵ -cyclase, isopentenyl pyrophosphate isomerase and β -carotene hydroxylase as well as vectors containing the same and hosts transformed with the vectors. Methods for controlling the ratio of various carotenoids in a host and for the production of novel carotenoid pigments. The present invention also provides a method for screening for eukaryotic genes encoding carotenoid biosynthesis, and for modifying the disclosed enzymes.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| AL | Albania | ES | Spain | LS | Lesotho | SI | Slovenia |
|----|--------------------------|-----|---------------------|----|-----------------------|-----|--------------------------|
| AM | Armenia | FI | Finland | LT | Lithuania | SK | Slovakia |
| AT | Austria | FR | France | LU | Luxembourg | SN | Senegal |
| ΑU | Australia | GA | Gabon | LV | Latvia | SZ | Swaziland |
| AZ | Azerbaijan | GB | United Kingdom | MC | Monaco | TD | Chad |
| BA | Bosnia and Herzegovina | GE | Georgia | MD | Republic of Moldova | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagascar | TJ | Tajikistan |
| BE | Belgium | GN | Guinea | MK | The former Yugoslav | TM | Turkmenistan |
| BF | Burkina Faso | GR | Greece | | Republic of Macedonia | TR | Turkey |
| BG | Bulgaria | HU | Hungary | ML | Mali | TT | Trinidad and Tobago |
| BJ | Benin _. | 16 | Ireland | MN | Mongolia | UA | Ukraine |
| BR | Brazil | IL | Israel | MR | Mauritania | UG | Uganda |
| BY | Belarus | IS | Iceland | MW | Malawi | US | United States of America |
| CA | Canada | IT | Italy | MX | Mexico | UZ | Uzbekistan |
| CF | Central African Republic | JP | Japan | NE | Niger | VN | Viet Nam |
| CG | Congo | KE | Kenya | NL | Netherlands | YU | Yugoslavia |
| CH | Switzerland | KG | Kyrgyzstan | NO | Norway | zw | Zimbabwe |
| CI | Côte d'Ivoire | KP | Democratic People's | NZ | New Zealand | 2,, | Ziniozowe |
| CM | Cameroon | | Republic of Korea | PL | Poland | | |
| CN | China | KR | Republic of Korea | PT | Portugal | | |
| CU | Cuba | KZ. | Kazakstan | RO | Romania | | |
| CZ | Czech Republic | LC | Saint Lucia | RU | Russian Federation | | |
| DE | Germany | LI | Liechtenstein | SD | Sudan | | |
| DK | Denmark | LK | Sri Lanka | SE | Sweden | | |
| EE | Estonia | LR | Liberia | SG | Singapore | | |



BACKGROUND OF THE INVENTION

Field of the Invention

5

10

The present invention describes nucleic acid sequences for eukaryotic genes encoding ϵ lycopene ϵ -cyclase (also known as ϵ -cyclase and ϵ lycopene cyclase), isopentenyl pyrophosphate isomerase (IPP) and β -carotene hydroxylase as well as vectors containing the same and hosts transformed with said vectors. The present invention also provides methods for augmenting the accumulation of carotenoids, changing the composition of the carotenoids, and producing novel and rare carotenoids. The present invention provides methods for controlling the ratio or relative amounts of various carotenoids in a host. The invention also relates to modified lycopene ϵ -cyclase, IPP isomerase and β -carotene hydroxylase. Additionally, the present invention provides a method for screening for genes and cDNAs encoding enzymes of carotenoid biosynthesis and metabolism.

15

20

Background of the Invention

Carotenoid pigments with cyclic endgroups are essential components of the photosynthetic apparatus in oxygenic photosynthetic organisms (e.g., cyanobacteria, algae and plants; Goodwin, 1980). The symmetrical bicyclic yellow carotenoid pigment βcarotene (or, in rare cases, the asymmetrical bicyclic \alpha-carotene) is intimately associated with the photosynthetic reaction centers and plays a vital role in protecting against potentially lethal photooxidative damage (Koyama, 1991). β-carotene and other carotenoids derived from it or from α-carotene also serve as light-harvesting pigments (Siefermann-Harms, 1987), are involved in the thermal dissipation of excess light energy captured by the lightharvesting antenna (Demmig-Adams & Adams, 1992), provide substrate for the biosynthesis of the plant growth regulator abscisic acid (Rock & Zeevaart, 1991; Parry & Horgan, 1991), and are precursors of vitamin A in human and animal diets (Krinsky, 1987). Plants also exploit carotenoids as coloring agents in flowers and fruits to attract pollinators and agents of seed dispersal (Goodwin, 1980). The color provided by carotenoids is also of agronomic value in a number of important crops. Carotenoids are currently harvested from a variety of organisms, including plants, algae, yeasts, cyanobacteria and bacteria, for use as pigments in food and feed.

30

The probable pathway for formation of cyclic carotenoids in plants, algae and cyanobacteria is illustrated in Figure 1. Two types of cyclic endgroups or rings are commonly found in higher plant carotenoids, these are referred to as the β (beta) and ϵ (epsilon) rings (Fig. 3). The precursor acyclic endgroup (no ring structure) is referred to as the Ψ (psi) endgroup. The β and ϵ endgroups differ only in the position of the double bond in the ring. Carotenoids with two β rings are ubiquitous, and those with one β and one ϵ ring are common, but carotenoids with two ϵ rings are uncommon. β -carotene (Fig. 1) has two β -endgroups and is a symmetrical compound that is the precursor of a number of other important plant carotenoids such as zeaxanthin and violaxanthin (Fig. 2).

10

15

20

25

5

Genes encoding enzymes of carotenoid biosynthesis have previously been isolated from a variety of sources including bacteria (Armstrong et al., 1989, Mol. Gen. Genet. 216, 254-268; Misawa et al., 1990, J. Bacteriol., 172, 6704-12), fungi (Schmidhauser et al., 1990, Mol. Cell. Biol. 10, 5064-70), cyanobacteria (Chamovitz et al., 1990, Z. Naturforsch, 45c, 482-86; Cunningham et al., 1994) and higher plants (Bartley et al., Proc. Natl. Acad. Sci USA 88, 6532-36; Martinez-Ferez & Vioque, 1992, Plant Mol. Biol. 18, 981-83). Many of the isolated enzymes show a great diversity in structure, function and inhibitory properties between sources. For example, phytoene desaturases from the cyanobacterium Synechococcus and from higher plants and green algae carry out a two-step desaturation to yield ζ-carotene as a reaction product. In plants and cyanobacteria a second enzyme (ζcarotene desaturase), similar in amino acid sequence to the phytoene desaturase, catalyzes two additional desaturations to yield lycopene. In contrast, a single desaturase enzyme from Erwinia herbicola and from other bacteria introduces all four double bonds required to form lycopene. The Erwinia and other bacterial desaturases bear little amino acid sequence similarity to the plant and cyanobacterial desaturase enzymes, and are thought to be of unrelated ancestry. Therefore, even with a gene in hand from one source, it may be difficult to identify a gene encoding an enzyme of similar function in another organism. In particular, the sequence similarity between certain of the prokaryotic and eukaryotic genes encoding enzymes of carotenoid biosynthesis is quite low.

30

Further, the mechanism of gene expression in prokaryotes and eukaryotes appears to differ sufficiently such that one cannot expect that an isolated eukaryotic gene will be properly expressed in a prokaryotic host.

The difficulties in isolating genes encoding enzymes with similar functions is exemplified by recent efforts to isolate the gene encoding the enzyme that catalyzes the formation of β -carotene from the acyclic precursor lycopene. Although a gene encoding an enzyme with this function had been isolated from a bacterium, it had not been isolated from any photosynthetic procaryote or from any eukaryotic organism. The isolation and characterization of the enzyme catalyzing formation of β -carotene in the cyanobacterium Synechococcus PCC7942 was described by the present inventors and others (Cunningham et al., 1993 and 1994). The amino acid sequence similarity of the cyanobacterial enzyme to the various bacterial lycopene β -cyclases is so low (ca. 18-25% overall; Cunningham et al., 1994) that there is much uncertainty as to whether they share a common ancestry or, instead, represent an example of convergent evolution.

The need remains for the isolation of eukaryotic and prokaryotic genes and cDNAs encoding polypeptides involved in the carotenoid biosynthetic pathway, including those encoding a lycopene ϵ -cyclase, IPP isomerase and β -carotene hydroxylase. There remains a need for methods to enhance the production of carotenoids, to alter the composition of carotenoids, and to reduce or eliminate carotenoid production. There also remains a need in the art for methods for screening for genes and cDNAs encoding enzymes of carotenoid biosynthesis and metabolism.

SUMMARY OF THE INVENTION

20

25

30

5

10

15

Accordingly, a first object of this invention is to provide purified and/or isolated nucleic acids which encode enzymes involved in carotenoid biosynthesis; in particular, lycopene ε-cyclase, IPP isomerase and β-carotene hydroxylase.

A second object of this invention is to provide purified and/or isolated nucleic acids which encode enzymes which produce novel or uncommon carotenoids.

A third object of the present invention is to provide vectors containing said genes.

A fourth object of the present invention is to provide hosts transformed with said vectors.

Another object of the present invention is to provide hosts which accumulate novel or uncommon carotenoids or which accumulate greater amounts of specific or total carotenoids.

Another object of the present invention is to provide hosts with inhibited and/or altered carotenoid production.

10

15

20

25

30

Another object of this invention is to secure the expression of eukaryotic carotenoidrelated genes in a recombinant prokaryotic host.

Yet another object of the present invention is to provide a method for screening for eukaryotic and prokaryotic genes and cDNAs which encode enzymes involved in carotenoid biosynthesis and metabolism.

An additional object of the invention is to provide a method for manipulating carotenoid biosynthesis in photosynthetic organisms by inhibiting the synthesis of certain enzymatic products to cause accumulation of precursor compounds.

Another object of the invention is to provide modified lycopene ϵ -cyclase, IPP isomerase and β -carotene hydroxylase.

These and other objects of the present invention have been realized by the present inventors as described below.

A subject of the present invention is an isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ε-cyclase, IPP isomerase or β-carotene hydroxylase enzyme activity and having the amino acid sequence of SEQ ID NOS: 2, 4, 14-21 or 23-27.

The invention also includes vectors which comprise any of the nucleic acid sequences listed above, and host cells transformed with such vectors.

Another subject of the present invention is a method of producing or enhancing the production of a carotenoid in a host cell, comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence to produce the protein.

Yet another subject of the present invention is a method of modifying the production of carotenoids in a host cell, the method comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which produces an RNA and/or encodes for a protein which modifies lycopene ε-cyclase, IPP isomerase or β-carotene hydroxylase enzyme activity, relative to an untransformed host cell, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence in the host cell to modify the production of the carotenoids in the host cell, relative to the untransformed host cell.

10 .

15

20

25

30



The present invention also includes a method of expressing, in a host cell, a heterologous nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity, the method comprising inserting into the host cell a vector comprising the heterologous nucleic acid sequence, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence.

Also included is a method of expressing, in a host cell, a heterologous nucleic acid sequence which encodes for a protein which modifies lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity in the host cell, relative to an untransformed host cell, the method comprising inserting into the host cell a vector comprising the heterologous nucleic acid sequence, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence.

Another subject of the present invention is a method for screening for genes and cDNAs which encode enzymes involved in carotenoid biosynthesis and metabolism.

BRIEF DESCRIPTION OF THE DRAWINGS

A more complete appreciation of the invention and many of the attendant advantages thereof will be readily obtained as the same becomes better understood by reference to the following detailed description when considered in connection with the accompanying drawings, wherein:

Figure 1 is a schematic representation of the putative pathway of β -carotene biosynthesis in cyanobacteria, algae and plants. The enzymes catalyzing various steps are indicated at the left. Target sites of the bleaching herbicides NFZ and MPTA are also indicated at the left. Abbreviations: DMAPP, dimethylallyl pyrophosphate; FPP, farnesyl pyrophosphate; GGPP, geranylgeranyl pyrophosphate; GPP, geranyl pyrophosphate; IPP, isopentenyl pyrophosphate; LCY, lycopene cyclase; MVA, mevalonic acid; MPTA, 2-(4-methylphenoxy)triethylamine hydrochloride; NFZ, norflurazon; PDS, phytoene desaturase; PSY, phytoene synthase; ZDS, ζ -carotene desaturase; PPPP, prephytoene pyrophosphate.

Figure 2 depicts possible routes of synthesis of cyclic carotenoids and common plant and algal xanthophylls (oxycarotenolds) from neurosporene. Demonstrated activities of the β - and ϵ -cyclase enzymes of A. thaliana are indicated by bold arrows labelled with β or ϵ respectively. A bar below the arrow leading to ϵ -carotene indicates that the enzymatic

10

15

20

25

30



activity was examined but no product was detected. The steps marked by an arrow with a dotted line have not been specifically examined. Conventional numbering of the carbon atoms is given for neurosporene and α -carotene. Inverted triangles (∇) mark positions of the double bonds introduced as a consequence of the desaturation reactions.

Figure 3 depicts the carotene endgroups which are found in plants.

Figure 4 is a DNA sequence and the predicted amino acid sequence of a lycopene ε-cyclase cDNA isolated from *A. thaliana* (SEQ ID NOS: 1 and 2). These sequences were deposited under Genbank accession number U50738. This cDNA is incorporated into the plasmid pATeps.

Figure 5 is a DNA sequence encoding the β -carotene hydroxylase isolated from A. thaliana (SEQ ID NO: 3). This cDNA is incorporated into the plasmid pATOHB.

Figure 6 is an alignment of the predicted amino acid sequences of A. thaliana β-carotene hydroxylase (SEQ ID NO: 4) with those of the bacterial β-carotene hydroxylase enzymes from Alicalgenes sp. (SEQ ID NO: 5) (Genbank D58422), Erwinia herbicola Eho10 (SEQ ID NO.: 6) (GenBank M872280), Erwinia uredovora (SEQ ID NO.: 7) (GenBank D90087) and Agrobacterium aurianticum (SEQ ID NO.: 8) (GenBank D58420). A consensus sequence is also shown. All five genes are identical where a capital letter appears in the consensus. A lowercase letter indicates that three of five, including A. thaliana, have the identical residue. TM; transmembrane.

Figure 7 is a DNA sequence of a cDNA encoding an IPP isomerase isolated from A. thaliana (SEQ ID NO: 9). This cDNA is incorporated into the plasmid pATDP5.

Figure 8 is a DNA sequence of a second cDNA encoding another IPP isomerase isolated from A. thaliana (SEQ ID NO: 10). This cDNA is incorporated into the plasmid pATDP7.

Figure 9 is a DNA sequence of a cDNA encoding an IPP isomerase isolated from *Haematococcus pluvialis* (SEQ ID NO: 11). This cDNA is incorporated into the plasmid pHP04.

Figure 10 is a DNA sequence of a second cDNA encoding another IPP isomerase isolated from *Haematococcus pluvialis* (SEQ ID NO: 12). This cDNA is incorporated into the plasmid pHP05.

Figure 11 is an alignment of the amino acid sequences predicted by IPP isomerase cDNAs isolated from A. thaliana (SEQ ID NO.: 16 and 18), H. pluvialis (SEQ ID NOS.: 14

WO 99/63055 PCT/US99/12121

and 15), Clarkia breweri (SEQ ID NO.: 17) (See, Blanc & Pichersky, Plant Physiol. (1995) 108:855; Genbank accession no. X82627) and Saccharomyces cerevisiae (SEQ ID NO.: 19) (Genbank accession no. J05090).

Figure 12 is a DNA sequence of the cDNA encoding an IPP isomerase isolated from *Tagetes erecta* (marigold; SEQ ID NO: 13). This cDNA is incorporated into the plasmid pPMDP1. xxx's denote a region not originally sequenced. Figure 21A shows the complete marigold sequence.

Figure 13 is an alignment of the consensus sequence of four plant β -cyclases (SEQ ID NO.: 20) with the A. thaliana lycopene ϵ -cyclase (SEQ ID NO.: 21). A capital letter in the plant β consensus is used where all four β -cyclase genes predict the same amino acid residue in this position. A small letter indicates that an identical residue was found in three of the four. Dashes indicate that the amino acid residue was not conserved and dots in the sequence denote a gap. A consensus for the aligned sequences is given, in capital letters below the alignment, where the β - and ϵ -cyclases have the same amino acid residue. Arrows indicate some of the conserved amino acids that will be used as junction sites for construction of chimeric cyclases with novel enzymatic activities. Several regions of interest including a sequence signature indicative of a dinucleotide-binding motif and two predicted transmembrane (TM) helical regions are indicated below the alignment and are underlined.

Figure 14 shows the nucleotide (SEQ ID NO:22) and amino acid sequences (SEQ ID NO:23) of the *Adonis palaestina* (pheasant's eye) ϵ -cyclase cDNA #5.

Figure 15A shows the nucleotide (SEQ ID NO:24) and amino acid sequences (SEQ ID NO:25) of a potato ε-cyclase cDNA. Figure 15B shows the amino acid sequence (SEQ ID NO:26) of a chimeric lettuce/potato lycopene ε-cyclase. Amino acids in lower case are from the lettuce cDNA and those in upper case are from the potato cDNA. The product of this chimeric cDNA has e-cyclase activity and converts lycopene to the monocyclic δ-carotene.

Figure 16 shows a comparison between the amino acid sequences of the *Arabidopsis* ϵ -cyclase (SEQ ID NO:27) and the potato ϵ -cyclase (SEQ ID NO:25).

Figure 17A shows the nucleotide sequence of the *Adonis palaestina* Ipi1 (SEQ ID NO:28) and Figure 17B shows the nucleotide sequence of the *Adonis palaestina* Ipi2 (SEQ ID NO: 29).

5

10

15

20

25

10

15

20

25

30

Figure 18A shows the nucleotide sequence of the *Haematoccus pluvialis* Ipi1 (SEQ ID NO:11) and Figure 18B shows the nucleotide sequence of the *Haematoccus pluvialis* Ipi2 (SEQ ID NO:30).

Figure 19A shows the nucleotide sequence of the *Lactuca sativa (romaine lettuce)* Ipi1 (SEQ ID NO:31) and Figure 19B shows the nucleotide sequence of the *Lactuca sativa* Ipi2 (SEQ ID NO: 32).

Figure 20 shows the nucleotide sequence of the *Chlamydomonas reinhardtii* Ipi1 (SEQ ID NO:33).

Figure 21A shows the nucleotide sequence of the *Tagetes erecta* (marigold) Ipi1 (SEQ ID NO:34) and Figure 21B shows the nucleotide sequence of the *Oryza sativa* (rice) Ipi1 (SEQ ID NO:35).

Figure 22 shows a amino acid sequence alignment of various plant and green algal isopentenyl isomerases (IPI) (SEQ ID NOS:16, 36-45).

Figure 23 shows a comparison between *Adonis palaestina* ϵ -cyclase cDNA #3 and cDNA #5 nucleotide sequences.

Figure 24 shows a comparison between *Adonis palaestina* ϵ -cyclase cDNA #3 and cDNA #5 predicted amino acid sequences.

Figure 25 shows a sequence alignment of various plant β - and ϵ -cyclases. Those sequences outlined in grey denote identical sequences among the ϵ -cyclases. Those sequences outlined in black denote identical sequences among both the β - and ϵ -cyclases.

Figure 26 shows a sequence alignment of the plant ϵ -cyclases from Figure 25. Those sequences outlined in black denote identical sequences among the ϵ -cyclases.

Figure 27 is a dendrogram or "tree" illustrating the degree of amino acid sequence similarity for various lycopene β - and ϵ -cyclases.

Figure 28 shows a comparison between Arabidopsis ϵ -cyclase and lettuce ϵ -cyclase predicted amino acid sequences.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention includes an isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ε-cyclase, IPP isomerase or β-carotene hydroxylase enzyme activity and having the amino acid sequence of SEQ ID NOS: 2, 4, 14-21, 23 or 25-27. Nucleic acids encoding lycopene ε-cyclase, β-carotene hydroxylase and IPP

10

15

20

25

30

055

isomerases have been isolated from several genetically distant sources.

The present inventors have isolated nucleic acids encoding the enzyme IPP isomerase, which catalyzes the reversible conversion of isopentenyl pyrophosphate (IPP) to dimethylallyl pyrophosphate (DMAPP). IPP isomerase cDNAs were isolated from the plants A. thaliana, Tagetes erecta (marigold), Adonis palaestina (pheasant's eye), Lactuca sativa (romaine lettuce) and from the green algae H. pluvialis and Chlamydomonas reinhardtii.

Alignments of the amino acid sequences predicted by some of these cDNAs are shown in Figures 12 and 22. Plasmids containing some of these cDNAs were deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville MD 20852 on March 4, 1996 under ATCC accession numbers 98000 (pHP05 - H. pluvialis); 98001 (pMDP1 - marigold); 98002 (pATDP7 - A. thaliana) and 98004 (pHP04 - H. pluvialis).

The present inventors have also isolated nucleic acids encoding the enzyme β-carotene hydroxylase, which is responsible for hydroxylating the β-endgroup in carotenoids. The nucleic acid of the present invention is shown in SEQ ID NO: 3 and Figure 5. The full length cDNA product hydroxylates both end groups of β-carotene as do products of cDNAs which encode proteins truncated by up to 50 amino acids from the N-terminus. Products of genes which encode proteins truncated between about 60-110 amino acids from the N-terminus preferentially hydroxylate only one ring. A plasmid containing this gene was deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville MD 20852 on March 4, 1996 under ATCC accession number 98003 (pATOHB - A. thaliana).

The present inventors have also isolated nucleic acids encoding the enzyme lycopene ϵ -cyclase, which is responsible for the formation of ϵ -endgroups in carotenoids. The A. thaliane ϵ -cyclase adds an ϵ ring to only one end of the symmetrical lycopene while the related β -cyclase adds a ring at both ends. The A. thaliana cDNA of the present invention is shown in Figure 4 and SEQ ID NO: 1. A plasmid containing this gene was deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville MD 20852 on March 4, 1996 under ATCC accession number 98005 (pATeps - A. thaliana).

In addition, lycopene ϵ -cyclases have been identified in lettuce and in *Adonis* palaestina (cDNA #5) which encode enzymes that convert lycopene to the bicyclic ϵ -carotene (ϵ , ϵ -carotene). An additional cDNA from *Adonis palaestina* (cDNA #3) encodes a lycopene ϵ -cyclase which converts lycopene into δ -carotene (ϵ , ψ -carotene) and differs from the lycopene ϵ -cyclase which forms bicyclic ϵ -carotene (ϵ , ϵ -carotene) by only 5 amino acids.

10

15

20

25

30

One or more of these amino acids may be modified by alteration of the nucleotide sequence in the #5 cDNA to obtain an enzyme which forms the bicyclic ϵ,ϵ -carotene. The sequences of the *Adonis palaestina* and *Arabidopsis thaliana* ϵ -cyclases have about 70% nucleotide identity and about 72% amino acid identity.

Initial experiments by the inventors with chimeric genes indicated that the part of the ϵ -cyclase which is responsible for adding 2ϵ rings to form ϵ, ϵ -carotene is the carboxy terminal portion of the gene. The lettuce ϵ -cyclase adds two ϵ rings to form ϵ, ϵ -carotene. A DNA encoding a partial potato ϵ -cyclase (missing its amino terminal portion), when combined with an amino terminal region from the lettuce ϵ -cyclase gene, produces a monocyclic δ -carotene (ϵ, ψ -carotene). With the discovery of the differences between the Adonis palaestina clone #3 and clone #5, the specific amino acids responsible for the addition of an extra ϵ ring have been identified (Figure 24). Specifically, amino acid 55 is Thr in clone #3 and Ser in clone #5, amino acid 210 is Asn in clone #3 and Asp in clone #5, amino acid 231 is Asp in clone #3 and Glu in clone #5, amino acid 352 is Ile in clone #3 and Val in clone #5, and amino acid 524 is Lys in clone #3 and Arg in clone #5. It can be appreciated that these changes are quite conservative, as only one change, at amino acid 210, changes the charge of the protein.

Thus, it is clear that the nucleic acids of the invention encoding the enzymes as presently disclosed may be altered to increase a particularly desirable property of the enzyme, to change a property of the enzyme, or to diminish an undesirable property of the enzyme. Such modifications can be by deletion, substitution, or insertion of one or more amino acids, and can be performed by routine enzymatic manipulation of the nucleic acid encoding the enzyme (such as by restriction enzyme digestion, removal of nucleotides by mung bean nuclease or *Bal*31, insertion of nucleotides by Klenow fragment, and by religation of the ends), by site-directed mutagenesis, or may be accidental, such as by low fidelity PCR or those obtained through mutations in hosts that are producers of the enzymes. These techniques as well as other suitable techniques are well known in the art.

Mutations can be made in the nucleic acids of the invention such that a particular codon is changed to a codon which codes for a different amino acid. Such a mutation is generally made by making the fewest nucleotide changes possible. A substitution mutation of this sort can be made to change an amino acid in the resulting protein in a non-conservative manner (i.e., by changing the codon from an amino acid belonging to a grouping

15

20

25

30

of amino acids having a particular size or characteristic to an amino acid belonging to another grouping) or in a conservative manner (i.e., by changing the codon from an amino acid belonging to a grouping of amino acids having a particular size or characteristic to an amino acid belonging to the same grouping). Such a conservative change generally leads to less change in the structure and function of the resulting protein. A non-conservative change is more likely to alter the structure, activity or function of the resulting protein. The present invention should be considered to include sequences containing conservative changes which do not significantly alter the activity or binding characteristics of the resulting protein.

The following is one example of various groupings of amino acids:

Amino acids with nonpolar R groups: Alanine, Valine, Leucine, Isoleucine, Proline, Phenylalanine, Tryptophan and Methionine.

Amino acids with uncharged polar R groups: Glycine, Serine, Threonine, Cysteine, Tyrosine, Asparagine and Glutamine.

Amino acids with charged polar R groups (negatively charged at Ph 6.0): Aspartic acid and Glutamic acid.

Basic amino acids (positively charged at pH 6.0): Lysine, Arginine and Histidine.

Another grouping may be those amino acids with phenyl groups: Phenylalanine, Tryptophan and Tyrosine.

Another grouping may be according to molecular weight (i.e., size of R groups). Particularly preferred substitutions are:

- Lys for Arg and vice versa such that a positive charge may be maintained;
- Glu for Asp and vice versa such that a negative charge may be maintained;
- Ser for Thr such that a free -OH can be maintained; and
- Gln for Asn such that a free NH₂ can be maintained.

Amino acid substitutions may also be introduced to substitute an amino acid with a particularly preferable property. For example, a Cys may be introduced to provide a potential site for disulfide bridges with another Cys. A His may be introduced as a particularly "catalytic" site (i.e., His can act as an acid or base and is the most common amino acid in biochemical catalysis). Pro may be introduced because of its particularly planar structure, which induces β-turns in the protein's structure.

It is clear that certain modifications of SEQ ID NOS: 2, 4, 14-21, 23 or 25-27 can take place without destroying the activity of the enzyme. It is noted especially that truncated

versions of the nucleic acids of the invention are functional. For example, several amino acids (from 1 to about 120) can be deleted from the N-terminus of the lycopene ε-cyclases of the invention, and a functional protein can still be produced. This fact is made especially clear from Figure 25, which shows a sequence alignment of several plant ε-cyclases. As can be seen from Figure 25, there is an enormous amount of sequence disparity between amino acid sequences 2 to about 50-70 (depending on the particular sequence, since gaps are present). There is less, but also a substantial amount of, sequence dissimilarity between about 50-70 to about 90-120 (depending on the particular sequence). Thereafter, the sequences are fairly conserved, except for small pockets of dissimilarity between about 275-295 to about 285-305 (depending on the particular sequence), and between about 395-415 to about 410-430 (depending on the particular sequence).

The present inventors have found that the amount of the 5' region present in the nucleic acids of the invention can alter the activity of the enzyme. Instead of diminishing activity, truncating the 5' region of the nucleic acids of the invention may result in an enzyme with a different specificity. Thus, the present invention relates to nucleic acids and enzymes encoded thereby which are truncated to within 0-50, preferably 0-25, codons of the 5' initiation codon of their prokaryotic counterparts as determined by alignment maps as discussed below.

For example, when the cDNA encoding A. thaliana β -carotene hydroxylase was truncated, the resulting enzyme catalyzed the formation of β -cryptoxanthin as the major product and zeaxanthin as minor product; in contrast to its normal production of zeaxanthin.

The present invention is intended to include those nucleic acid and amino acid sequences in which substitutions, deletions, additions or other modifications have taken place, as compared to SEQ ID NOS: 2, 4, 14-21, 23 or 25-27, without destroying the activity of the enzyme. Preferably, the substitutions, deletions, additions or other modifications take place at the 5' end, or any other of those positions which already show dissimilarity between any of the presently disclosed amino acid sequences (see also Figure 25) or other amino acid sequences which are known in the art and which encode the same enzyme (i.e., lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase).

In each case, nucleic acid and amino acid sequence similarity and identity is measured using sequence analysis software, for example, the Sequence Analysis, Gap, or BestFit software packages of the Genetics Computer Group (University of Wisconsin Biotechnology

5

10

15

20

25

Center, 1710 University Avenue, Madison, Wisconsin 53705), MEGAlign (DNAStar, Inc., 1228 S. Park St., Madison, Wisconsin 53715), or MacVector (Oxford Molecular Group, 2105 S. Bascom Avenue, Suite 200, Campbell, California 95008). Such software uses algorithms to match similar sequences by assigning degrees of identity to various substitutions, deletions, and other modifications, and includes detailed instructions as to useful parameters, etc., such that those of routine skill in the art can easily compare sequence similarities and identities. An example of a useful algorithm in this regard is the algorithm of Needleman and Wunsch, which is used in the Gap program discussed above. This program finds the alignment of two complete sequences that maximizes the number of matches and minimizes the number of gaps. Another useful algorithm is the algorithm of Smith and Waterman, which is used in the BestFit program discussed above. This program creates an optimal alignment of the best segment of similarity between two sequences. Optimal alignments are found by inserting gaps to maximize the number of matches using the local homology algorithm of Smith and Waterman.

15

10

5

Conservative (i.e. similar) substitutions typically include substitutions within the following groups: glycine and alanine; valine, isoleucine and leucine; aspartic acid, glutamic acid, asparagine and glutamine; serine and threonine; lysine and arginine; and phenylalanine and tyrosine. Substitutions may also be made on the basis of conserved hydrophobicity or hydrophilicity (see Kyte and Doolittle, *J. Mol. Biol.* 157: 105-132 (1982)), or on the basis of the ability to assume similar polypeptide secondary structure (see Chou and Fasman, *Adv. Enzymol.* 47: 45-148 (1978)).

20

25

If comparison is made between nucleotide sequences, preferably the length of comparison sequences is at least 50 nucleotides, more preferably at least 60 nucleotides, at least 75 nucleotides or at least 100 nucleotides. It is most preferred if comparison is made between the nucleic acid sequences encoding the enzyme coding regions necessary for enzyme activity. If comparison is made between amino acid sequences, preferably the length of comparison is at least 20 amino acids, more preferably at least 30 amino acids, at least 40 amino acids or at least 50 amino acids. It is most preferred if comparison is made between the amino acid sequences in the enzyme coding regions necessary for enzyme activity.

30

It should be appreciated that also within the scope of the present invention are nucleic acid sequences encoding lycopene ε-cyclases, IPP isomerases and β-carotene hydroxylases

10

15

20

25

30

which code for enzymes having the same amino acid sequence as SEQ ID NOS: 2, 4, 14-21, 23 or 25-27, but which are degenerate to the nucleic acids specifically disclosed herein.

The amino acid residues described herein are preferred to be in the "L" isomeric form. However, residues in the "D" isomeric form can be substituted for any L-amino acid residue, as long as the desired functional property of immunoglobulin-binding is retained by the polypeptide.

In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Sambrook et al, "Molecular Cloning: A Laboratory Manual" (1989); "Current Protocols in Molecular Biology" Volumes I-III [Ausubel, R. M., ed. (1994)]; "Cell Biology: A Laboratory Handbook" Volumes I-III [J. E. Celis, ed. (1994))]; "Current Protocols in Immunology" Volumes I-III [Coligan, J. E., ed. (1994)]; "Oligonucleotide Synthesis" (M.J. Gait ed. 1984); "Nucleic Acid Hybridization" [B.D. Hames & S.J. Higgins eds. (1985)]; "Transcription And Translation" [B.D. Hames & S.J. Higgins, eds. (1984)]; "Animal Cell Culture" [R.I. Freshney, ed. (1986)]; "Immobilized Cells And Enzymes" [IRL Press, (1986)]; B. Perbal, "A Practical Guide To Molecular Cloning" (1984).

The present invention also includes vectors. Suitable vectors according to the present invention comprise a nucleic acid of the invention encoding an enzyme involved in carotenoid biosynthesis or metabolism and a suitable promoter for the host, and can be constructed using techniques well known in the art (for example Sambrook et al., Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; Ausubel et al., Current Protocols in Molecular Biology, Greene Publishing and Wiley Interscience, New York, 1991). Suitable vectors for eukaryotic expression in plants are described in Frey et al., Plant J. (1995) 8(5):693 and Misawa et al, 1994a; incorporated herein by reference. Suitable vectors for prokaryotic expression include pACYC184, pUC119, and pBR322 (available from New England BioLabs, Bevery, MA) and pTrcHis (Invitrogen) and pET28 (Novagen) and derivatives thereof. The vectors of the present invention can additionally contain regulatory elements such as promoters, repressors, selectable markers such as antibiotic resistance genes, etc.

The nucleic acids encoding the carotenoid enzymes as described above, when cloned into a suitable expression vector, can be used to overexpress these enzymes in a plant

10

15

20

25

30



expression system or to inhibit the expression of these enzymes. For example, a vector containing the gene encoding lycopene ϵ -cyclase can be used to increase the amount of α -carotene and carotenoids derived from α -carotene (such as lutein and α -cryptoxanthin) in an organism and thereby alter the nutritional value, pharmacology and visual appearance value of the organism.

Therefore, the present invention includes a method of producing or enhancing the production of a carotenoid in a host cell, relative to an untransformed host cell, the method comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence to produce the protein.

The present invention also includes a method of modifying the production of carotenoids in a host cell, the method comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which produces an RNA and/or encodes for a protein which modifies lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity, relative to an untransformed host cell, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence in the host cell to modify the production of the carotenoids in the host cell, relative to the untransformed host cell.

The term "modifying the production" means that the amount of carotenoids produced in the host cell can be enhanced, reduced, or left the same, as compared to the untransformed host cell. In accordance with one embodiment of the present invention, the make-up of the carotenoids (i.e., the specific carotenoids produced) is changed vis a vis each other, and this change in make-up may result in either a net gain, net loss, or no net change in the total amount of carotenoids produced in the cell. In accordance with another embodiment of the present invention, the production or the biochemical activity of the carotenoids (or the enzymes which catalyze their formation) is enhanced by the insertion of an enzyme-encoding nucleic acid of the invention. In yet another embodiment of the invention, the production or the biochemical activity of the carotenoids (or the enzymes which catalyze their formation) may be reduced or inhibited by a number of different approaches available to those skilled in the art, including but not limited to such methodologies or approaches as anti-sense (e.g.,

10

15

20

25

30

Gray et al (1992) Plant Mol. Biol. 19:69-87), ribozymes (e.g., Wegener et al (1994) Mol. Gen. Genet. 245:465-470), co-suppression (e.g., Fray and Grierson (1993) Plant Mol. Biol. 22:589-602), targeted disruption of the gene (e.g., Schaefer et al. (1997) Plant J. 11:1195-1206), intracellular antibodies (e.g., Rondon and Marasco (1997) Ann. Rev. Microbiol. 51:257-283) or whatever other approaches rely on the knowledge or availability of the nucleic acid or amino acid sequences of the invention and/or portions thereof, to thereby reduce accumulation of carotenoids with ϵ rings and compounds derived from them (for ϵ -cyclase inhibition), or carotenoids with hydroxylated β rings and compounds derived from them (for β -hydroxylase inhibition), or, in the case if IPP isomerase, accumulation of any isoprenoid compound.

Preferably, at least a portion of the nucleic acid sequences used in the methods, vectors and host cells of the invention codes for an enzyme having an amino acid sequence which is at least 85% identical, preferably at least 90%, at least 95% or completely identical to SEQ ID NOS: 2, 4, 14-21, 23 or 25-27. Sequence identity is determined as noted above. Preferably, sequence additions, deletions or other modifications are made as indicated above, so as to not affect the function of the particular enzyme.

In a preferred embodiment, vectors are manufactured which contain a DNA encoding a eukaryotic IPP isomerase upstream of a DNA encoding a second eukaryotic carotenoid enzyme. The inventors have discovered that inclusion of an IPP isomerase gene increases the supply of substrate for the carotenoid pathway; thereby enhancing the production of carotenoid endproducts, as compared to a host cell which is not transformed with such a vector. This is apparent from the much deeper pigmentation in carotenoid-accumulating colonies of *E. coli* which also contain one of the aforementioned IPP isomerase genes when compared to colonies that lack this additional IPP isomerase gene. Similarly, a vector comprising an IPP isomerase gene can be used to enhance production of any secondary metabolite of dimethylallyl pyrophosphate and/or isopentenyl pyrophosphate (such as isoprenoids, steroids, carotenoids, etc.). The term "isoprenoid" is intended to mean any member of the class of naturally occurring compounds whose carbon skeletons are composed, in part or entirely, of isopentyl C₅ units. Preferably, the carbon skeleton is of an essential oil, a fragrance, a rubber, a carotenoid, or a therapeutic compound, such as paclitaxel.

A vector containing the cDNA encoding a lycopene ϵ -cyclase of the invention, preferably the lettuce lycopene ϵ -cyclase or Adonis ϵ -cyclase #5, can be used to increase the

10

15

20

25

30

amount of bicyclic ∈-carotene in an organism and thereby alter the nutritional value, pharmacology and visual appearance value of the organism. In addition, the transformed organism can be used in the formulation of therapeutic agents, for example in the treatment of cancer (see Mayne et al (1996) FASEB J. 10:690-701; Tsushima et al (1995) Biol. Pharm. Bull. 18:227-233).

An antisense strand of a nucleic acid of the invention can be inserted into a vector. For example, the lycopene ϵ -cyclase gene can be inserted into a vector and incorporated into the genomic DNA of a host, thereby inhibiting the synthesis of ϵ , β -carotenoids (lutein and α -carotene) and enhancing the synthesis of β , β -carotenoids (zeaxanthin and β -carotene).

The present invention also relates to novel enzymes which are encoded by the amino acid sequences of the invention, or portions thereof.

The present invention also relates to novel enzymes which can transform known carotenoids into novel or uncommon products. Currently ϵ -carotene (see Figure 2) and γ -carotene are commonly produced only in minor amounts. As described below, an enzyme can be produced which transforms lycopene to γ -carotene and lycopene to ϵ -carotene. With these products in hand, bulk synthesis of other carotenoids derived from them are possible. For example, ϵ -carotene can be hydroxylated to form lactucaxanthin, an isomer of lutein (one ϵ and one β ring) and zeaxanthin (two β rings) where both endgroups are, instead, ϵ rings.

In addition to novel enzymes produced by truncating the 5' region of known enzymes, as discussed above, novel enzymes which can participate in the formation of unusual carotenoids can be formed by replacing portions of one gene with an analogous sequence from a structurally related gene. For example, β -cyclase and ϵ -cyclase are structurally related (see Figure 13). By replacing a portion of β -lycopene cyclase with the analogous portion of ϵ -cyclase, an enzyme which produces γ -carotene will be produced (one β endgroup). Further, by replacing a portion of the lycopene ϵ -cyclase with the analogous portion of β -cyclase, an enzyme which produces ϵ -carotene will be produced (with some exceptions, such as the lettuce ϵ -cyclase, plant ϵ -cyclases normally produce a compound with one ϵ -endgroup, δ -carotene). Similarly, β -hydroxylase could be modified to produce enzymes of novel function by creation of hybrids with ϵ -hydroxylase.

Host systems according to the present invention can comprise any organism that already produces carotenoids or which has been genetically modified to produce carotenoids.

The IPP isomerase genes are more broadly applicable for enhancing production of any product dependent on DMAPP and/or IPP as a precursor.

Organisms which already produce carotenoids include plants, algae, some yeasts, fungi and cyanobacteria and other photosynthetic bacteria. Transformation of these hosts with vectors according to the present invention can be done using standard techniques such as those described in Misawa et al., (1990) supra; Hundle et al., (1993) supra; Hundle et al., (1991) supra; Misawa et al., (1991) supra; Sandmann et al., supra; and Schnurr et al., supra.

Transgenic organisms can be constructed which include the nucleic acid sequences of the present invention (Bird et al, 1991; Bramley et al, 1992; Misawa et al, 1994a; Misawa et al, 1994b; Cunningham et al, 1993). The incorporation of these sequences can allow the controlling of carotenoid biosynthesis, content, or composition in the host cell. These transgenic systems can be constructed to incorporate sequences which allow for the overexpression of the nucleic acids of the present invention. Transgenic systems can also be constructed containing antisense expression of the nucleic acid sequences of the present invention. Such antisense expression would result in the accumulation of the substrates of the substrates of the enzyme encoded by the sense strand.

A method for screening for eukaryotic genes which encode enzymes involved in carotenoid biosynthesis comprises transforming a prokaryotic host with a nucleic acid which may contain a eukaryotic or prokaryotic carotenoid biosynthetic gene; culturing said transformed host to obtain colonies; and screening for colonies exhibiting a different color than colonies of the untransformed host.

Suitable hosts include E. coli, cyanobacteria such as Synechococcus and Synechocystis, alga and plant cells. E. coli are preferred.

In a preferred embodiment, the above "color complementation" screening protocol can be enhanced by using mutants which are either (1) deficient in at least one carotenoid biosynthetic gene or (2) overexpress at least one carotenoid biosynthetic gene. In either case, such mutants will accumulate carotenoid precursors.

Prokaryotic and eukaryotic DNA or cDNA libraries can be screened in total for the presence of genes of carotenoid biosynthesis, metabolism and degradation. Preferred organisms to be screened include photosynthetic organisms.

5

10

15

20

25

10

15

20

25

30

E. coli can be transformed with these eukaryotic cDNA libraries using conventional methods such as those described in Sambrook et al, 1989 and according to protocols described by the vendors of the cloning vectors.

For example, the cDNA libraries in bacteriophage vectors such as lambdaZAP (Stratagene) or lambda ZIPLOX (Gibco BRL) can be excised en masse and used to transform *E.coli*.

Transformed *E. coli* can be cultured using conventional techniques. The culture broth preferably contains antibiotics to select and maintain plasmids. Suitable antibiotics include penicillin, ampicillin, chloramphenicol, etc. Culturing is typically conducted at 15-40°C, preferably at room temperature or slightly above (18-28°C), for 12 hours to 7 days.

Cultures are plated and the plates are screened visually for colonies with a different color than the colonies of the host $E.\ coli$ transformed with the empty plasmid cloning vector. For example, $E.\ coli$ transformed with the plasmid, pAC-BETA (described below), produce yellow colonies that accumulate β -carotene. After transformation with a cDNA library, colonies which contain a different hue than those formed by $E.\ coli/pAC$ -BETA would be expected to contain enzymes which modify the structure or accumulation of β -carotene. Similar $E.\ coli$ strains can be engineered which accumulate earlier products in carotenoid biosynthesis, such as lycopene, γ -carotene, etc.

Having generally described this invention, a further understanding can be obtained by reference to certain specific examples which are provided herein for purposes of illustration only and are not intended to be limiting unless otherwise specified.

EXAMPLE

I. Isolation of β-carotene hydroxylase

Plasmid Construction

An 8.6kb BgIII fragment containing the carotenoid biosynthetic genes of *Erwinia* herbicola was first cloned in the BamHI site of plasmid vector pACYC184 (chloramphenicol resistant), and then a 1.1kb BamHI fragment containing the *E. herbicola* β-carotene hydroxylase (*CrtZ*) was deleted. *E.coli* strains containing the resulting plasmid, pAC-BETA, accumulate β-carotene and form yellow colonies (Cunningham et al., 1994).

A full length cDNA encoding IPP isomerase of *Haematococcus pluvialis* (HP04) was first excised with *BamH*I and *Kpn*I from pBluescript SK-, and then ligated into the

corresponding sites of the pTrcHisA vector with high-level expression from the *trc* promoter (Invitrogen, Inc.). A fragment containing the IPP isomerase and *trc* promoter was subsequently excised with *EcoRV* and *KpnI*, treated with the Klenow fragment of DNA polymerase to produce blunt ends, and ligated in the Klenow-treated *HindIII* site of pAC-BETA. *E.coli* cells transformed with this new plasmid pAC-BETA-04 form orange colonies on LB plates (*vs.* yellow for those containing pAC-BETA) and cultures accumulate substantially more β-carotene (*ca.* two fold) than those that contain pAC-BETA.

Screening of an Arabidopsis cDNA Library

Several λ cDNA expression libraries of *Arabidopsis* were obtained from the *Arabidopsis* Biological Resource Center (Ohio State University, Columbus, OH) (Kieber et al., 1993). The λ cDNA libraries were excised *in vivo* using Stratagene's ExAssist SOLR system to produce a phagemid cDNA library wherein each phagemid contained also a gene conferring resistance to the antibiotic ampicillin.

E.coli strain DH10BZIP was chosen as the host cell for the screening and pigment production, although we have also used TOP10F' and XL1-Blue for this purpose. DH10B cells were transformed with plasmid pAC-BETA-04 and were plated on LB agar plates containing chloramphenicol at 50 µg/ml (from United States Biochemical Corporation). The phagemid Arabidopsis cDNA library was then introduced into DH10B cells already containing pAC-BETA-04. Transformed cells containing both pAC-BETA-04 and Arabidopsis cDNA library phagemids were selected on chloramphenicol plus ampicillin (150 μg/ml) agar plates. Maximum color development occurred after 3 to 7 days incubation at room temperature, and the rare bright yellow colonies were selected from a background of many thousands of orange colonies on each agar plate. Selected colonies were inoculated into 3 ml liquid LB medium containing ampicillin and chloramphenicol, and cultures were incubated at room temperature for 1-2 days, with shaking. Cells were then harvested by centrifugation and extracted with acetone in microfuge tubes. After centrifugation, the pigmented extract was spotted onto silica gel thin-layer chromatography (TLC) plates, and developed with a hexane:ether (1:1, by volume) mobile phases. B-carotene hydroxylaseencoding cDNAs were identified based on the appearance of a yellow pigment that comigrated with zeaxanthin on the TLC plates.

5

10

15

20

25



Subcloning and Sequencing

The plasmid containing the β-carotene hydroxylase cDNA was recovered and analyzed by standard procedures (Sambrook et al., 1989). The *Arabidopsis* β-carotene hydroxylase was sequenced completely on both strands on an automatic sequencer (Applied Biosystems, Model 373A, Version 2.0.1S). The cDNA insert of 0.95kb also was excised and ligated into the a pTrcHis vector. A *BgI*II restriction site within the cDNA was used to remove that portion of the cDNA that encodes the predicted polypeptide N terminal sequence region that is not also found in bacterial β-carotene hydroxylases (Figure 6). A BgIII-XhoI fragment was directionally cloned in BamHI-XhoI digested TrcHis vectors.

Pigment Analysis

5

10

15

20

25

30

A single colony was used to inoculate 50 ml of LB containing ampicillin and chloramphenicol in a 250-ml flask. Cultures were incubated at 28°C for 36 hours with gentle shaking, and then harvested at 5000 rpm in an SS-34 rotor. The cells were washed once with distilled H₂O and resuspended with 0.5 ml of water. The extraction procedures and HPLC were essentially as described previously (Cunningham et al, 1994).

II. Isolation and biochemical analysis of an Arabidopsis lycopene €-cyclase Plasmid Construction

Construction of plasmids pAC-LYC, pAC-NEUR, and pAC-ZETA is described in Cunningham et al., (1994). In brief, the appropriate carotenoid biosynthetic genes from *Erwinia herbicola*, *Rhodobacter capsulatus*, and *Synechococcus* sp. strain PCC7942 were cloned in the plasmid vector pACYC184 (New England BioLabs, Beverly, MA). Cultures of *E. coli* containing the plasmids pAC-ZETA, pAC-NEUR, and pAC-LYC, accumulate ζ-carotene, neurosporene, and lycopene, respectively. The plasmid pAC-ZETA was constructed as follows: an 8.6-kb BglII fragment containing the carotenoid biosynthetic genes of *E. herbicola* (GenBank M87280; Hundle et al., 1991) was obtained after partial digestion of plasmid pPL376 (Perry et al., 1986; Tuveson et al., 1986) and cloned in the BamHI site of pACYC184 to give the plasmid pAC-EHER. Deletion of adjacent 0.8- and 1.1-kb BamHI-BamHI fragments (deletion Z in Cunningham et al., 1994), and of a 1.1 kB Sall-Sall fragment (deletion X) served to remove most of the coding regions for the *E. herbicola* β-carotene hydroxylase (crtZ gene) and zeaxanthin glucosyltransferase (crtX gene), respectively. The

10

15

20

25

resulting plasmid, pAC-BETA, retains functional genes for geranylgeranyl pyrophosphate synthase (crtE), phytoene synthase (crtB), phytoene desaturase (crtI), and lycopene cyclase (crtY). Cells of E. coli containing this plasmid form yellow colonies and accumulate β -carotene. A plasmid containing both the lycopene ϵ - and β -cyclase cDNAs of A. thaliana was constructed by excising the ϵ -cyclase in clone y2 as a PvuI-PvuII fragment and ligating this piece in the SnaBI site of a plasmid (pSPORT 1 from GIBCO-BRL) that already contained the β -cyclase (Cunningham et al., 1996).

Organisms and Growth Conditions

E. coli strains TOP10 and TOP10 F' (obtained from Invitrogen Corporation, San Diego, CA) and XL1-Blue (Stratagene) were grown in Luria-Bertani (LB) medium (Sambrook et al., 1989) at 37°C in darkness on a platform shaker at 225 cycles per min. Media components were from Difco (yeast extract and tryptone) or Sigma (NaCl). Ampicillin at 150 μg/mL and/or chloramphenicol at 50 μg/mL (both from United States Biochemical Corporation) were used, as appropriate, for selection and maintenance of plasmids.

Mass Excision and Color Complementation Screening of an A. thaliana cDNA Library

A size-fractionated 1-2 kB cDNA library of A. thaliana in lambda ZAPII (Kieber et al., 1993) was obtained from the Arabidopsis Biological Resource Center at The Ohio State University (stock number CD4-14). Other size fractionated libraries were also obtained (stock numbers CD4-13, CD4-15, and CD4-16). An aliquot of each library was treated to cause a mass excision of the cDNAs and thereby produce a phagemid library according to the instructions provided by the supplier of the cloning vector (Stratagene; E. coli strain XL1-Blue and the helper phage R408 were used). The titre of the excised phagemid was determined and the library was introduced into a lycopene-accumulating strain of E. coli TOP10 F' (this strain contained the plasmid pAC-LYC) by incubation of the phagemid with the E. coli cells for 15 min at 37°C. Cells had been grown overnight at 30°C in LB medium supplemented with 2% (w/v) maltose and 10 mM MgSO₄ (final concentration), and harvested in 1.5 ml microfuge tubes at a setting of 3 on an Eppendorf microfuge (5415C) for 10 min. The pellets were resuspended in 10 mM MgSO₄ to a volume equal to one-half that of the

initial culture volume. Transformants were spread on large (150 mm diameter) LB agar petri plates containing antibiotics to provide for selection of cDNA clones (ampicillin) and maintenance of pAC-LYC (chloramphenicol). Approximately 10,000 colony forming units were spread on each plate. Petri plates were incubated at 37°C for 16 hr and then at room temperature for 2 to 7 days to allow maximum color development. Plates were screened visually with the aid of an illuminated 3x magnifier and a low power stage-dissecting microscope for the rare, pale pinkish-yellow to deep-yellow colonies that could be observed in the background of pink colonies. A colony color of yellow or pinkish-yellow was taken as presumptive evidence of a cyclization activity. These yellow colonies were collected with sterile toothpicks and used to inoculate 3ml of LB medium in culture tubes with overnight growth at 37°C and shaking at 225 cycles/min. Cultures were split into two aliquots in microfuge tubes and harvested by centrifugation at a setting of 5 in an Eppendorf 5415C microfuge. After discarding the liquid, one pellet was frozen for later purification of plasmid DNA. To the second pellet was added 1.5 ml EtOH, and the pellet was resuspended by vortex mixing, and extraction was allowed to proceed in the dark for 15-30 min with occasional remixing. Insoluble materials were pelleted by centrifugation at maximum speed for 10 min in a microfuge. Absorption spectra of the supernatant fluids were recorded from 350-550 nm with a Perkin Elmer lambda six spectrophotometer.

Analysis of isolated clones

20

5

10

15

Eight of the yellow colonies contained β -carotene indicating that a single gene product catalyzes both cyclizations required to form the two β endgroups of the symmetrical β -carotene from the symmetrical precursor lycopene. One of the yellow colonies contained a pigment with the spectrum characteristic of δ -carotene, a monocyclic carotenoid with a single ϵ endgroup. Unlike the β cyclase, this ϵ -cyclase appears unable to carry out a second cyclization at the other end of the molecule.

25

The observation that ϵ -cyclase is unable to form two cyclic ϵ -endgroups (e.g. the bicyclic ϵ -carotene) illuminates the mechanism by which plants can coordinate and control the flow of substrate into carotenoids derived from β -carotene versus those derived from α -carotene and also can prevent the formation of carotenoids with two ϵ endgroups.

30

The availability of the A. thaliana gene encoding the ϵ -cyclase enables the directed manipulation of plant and algal species for modification of carotenoid content and

PCT/US99/12121

composition. Through inactivation of the ϵ -cyclase, whether at the gene level by deletion of the gene or by insertional inactivation or by reduction of the amount of enzyme formed (by such as antisense technology), one may increase the formation of β -carotene and other pigments derived from it. Since vitamin A is derived only from carotenoids with β endgroups, an enhancement of the production of β -carotene versus α -carotene may enhance nutritional value of crop plants. Reduction of carotenoids with ϵ -endgroups may also be of value in modifying the color properties of crop plants and specific tissues of these plants. Alternatively, where production of α -carotene, or pigments such as lutein that are derived from α -carotene, is desirable, whether for the color properties, nutritional value or other reason, one may overexpress the ϵ -cyclase or express it in specific tissues. Wherever agronomic value of a crop is related to pigmentation provided by carotenoid pigments the directed manipulation of expression of the ϵ -cyclase gene and/or production of the enzyme may be of commercial value.

The predicted amino acid sequence of the A. thaliana ϵ -cyclase enzyme was determined. A comparison of the amino acid sequences of the β - and ϵ -cyclase enzymes of Arabidopsis thaliana (Fig. 13) as predicted by the DNA sequence of the respective cDNAs (Fig. 4 for the ϵ -cyclase cDNA sequence), indicates that these two enzymes have many regions of sequence similarity, but they are only about 37% identical overall at the amino acid level. The degree of sequence identity at the DNA base level, only about 50%, is sufficiently low such that we and others have been unable to detect this gene by hybridization using the β cyclase as a probe in DNA gel blot experiments.

REFERENCES

Each reference cited in this application and/or listed below is hereby incorporated by reference.

- 25 Bird et al, 1991 Biotechnology 9, 635-639.
 - Bishop et al., (1995) FEBS Lett. 367, 158-162.
 - Bramley, P.M. (1985) Adv. Lipid Res. 21, 243-279.
 - Bramley, P.M. (1992) Plant J. 2, 343-349.
 - Britton, G. (1988). Biosynthesis of carotenoids. In Plant Pigments, T.W. Goodwin, ed. (London: Academic Press), pp. 133-182.

30

5

10

15

15

20

25

30

Britton, G. (1979) Z. Naturforsch. Section C Biosci. 34, 979-985.

Britton, G. (1995) UV/Visible spectroscopy. In Carotenoids, Vol. IB: Spectroscopy,

G. Britton, S. Liaaen-Jensen, H.P. Pfander, eds. (Basel: Birkhauser Verlag), pp. 13-62.

Bouvier et al., (1994) Plant J. 6, 45-54.

Cunningham et al., (1985) Photochem. Photobiol. 42: 295-307.

Cunningham et al., (1993) FEBS Lett. 328, 130-138.

Cunningham et al., (1994) Plant Cell 6, 1107-1121.

Cunningham et al., (1996) Plant Cell 8, 1613-1626.

Davies, B.H. (1976). Carotenoids. In Chemistry and Biochemistry of Plant Pigments,

10 Vol. 2, T.W. Goodwin, ed (New York: Academic Press), pp. 38-165.

Del Sal et al., (1988). Nucl. Acids Res. 16, 9878.

Demmig-Adams & Adams, (1992) Ann. Rev. Plant Physiol. Mol. Biol. 43, 599-626.

Enzell & Back, (1995) Mass spectrometry. In Carotenoids, Vol. IB: Spectroscopy, G.

Britton, S. Liaaen-Jensen, H.P. Pfander, eds. (Basel: Birkhauser Verlag), pp. 261-320.

Frank & Cogdell (1993) Photochemistry and function of carotenoids in photosynthesis. In Carotenoids in Photosynthesis. A. Young and G. Britton, eds. (London: Chapman and Hall). pp. 253-326.

Goodwin, T.W. (1980). The Biochemistry of the Carotenoids. 2nd ed, Vol. 1 (London: Chapman and Hall.

Horvath et al., (1972) Phytochem. 11, 183-187.

Hugueney et al., (1995) Plant J. 8, 417-424.

Hundle et al., (1991) Photochem. Photobiol. 54, 89-93.

Jensen & Jensen, (1971) Methods Enzymol. 23, 586-602.

Kargl & Quackenbush, (1960) Archives Biochem. Biophys. 88, 59-63.

Kargl et al., (1960) Proc. Am. Hort. Soc. 75, 574-578.

Kieber et al., (1993) Cell 72, 427-441.

Koyama, Y. (1991) J. Photochem. Photobiol., B, 9, 265-80.

Krinsky, N.I. (1987) Medical uses of carotenoids. In Carotenoids, N.I. Krinsky, M.M. Mathews-Roth, and R.F. Taylor, eds. (New York: Plenum), pp. 195-206.

Kyte & Doolittle, (1982) J. Mol. Biol. 157, 105-132.

LaRossa & Schloss, (1984) J. Biol. Chem. 259, 8753-8757.

Misawa et al., (1994a) Plant J. 6, 481-489.

20

12.

Misawa et al., (1994b) J. Biochem, Tokyo, 116, 980-985.

Norris et al., (1995) Plant Cell 7, 2139-2149.

Pecker et al., (1996) Submitted to Plant Mol. Biol.

Perry et al., (1986) J. Bacteriol. 168, 607-612.

Persson & Argos, (1994) J. Mol. Biol. 237, 182-192.

Plumley & Schmidt, (1987) Proc. Nat. Acad. Sci. USA 83, 146-150.

Plumley & Schmidt, (1995) Plant Cell 7, 689-704.

Rossmann et al., (1974) Nature 250, 194-199.

Rock & Zeevaart (1991) Proc. Nat. Acad. Sci. USA 88, 7496-7499.

10 Rost et al., (1995) Protein Science 4, 521-533.

Sambrook et al., (1989) Molecular Cloning: A Laboratory Manual, 2nd edition (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Sancar, A. (1994) Biochemistry 33, 2-9.

Sander & Schneider, (1991) Proteins 9, 56-68.

15 Sandmann, G. (1994) Eur. J. Biochem. 223, 7-24.

Scolnik & Bartley, (1995) Plant Physiol. 108, 1342.

Siefermann-Harms, D. (1987) Physiol. Plant. 69, 561-568.

Spurgeon & Porter, (1980). Biosynthesis of carotenoids. In Biochemistry of

Isoprenoid Compounds, J.W. Porter, and S.L. Spurgeon, eds. (New York: Wiley), pp. 1-122.

Tomes, M.L. (1963) Bot. Gaz. 124, 180-185.

Tomes, M.L. (1967) Genetics 56, 227-232.

Tuveson et al., (1986) J. Bacteriol. 170, 4675-4680.

Van Beeumen et al., (1991) J. Biol. Chem. 266, 12921-12931.

Weedon & Moss, (1995) Structure and Nomenclature. In Carotenoids, Vol. IB:

25 Spectroscopy, G. Britton, S. Liaaen-Jensen, H.P. Pfander, eds. (Basel: Birkhauser Verlag), pp. 27-70.

Wierenga et al., (1986) J. Mol. Biol. 187, 101-107.

Zechmeister, L. (1962) Cis-Trans Isomeric Carotenoids, Vitamins A and Arylpolyenes. Springer-Verlag, Vienna.

Having now fully described the invention, it will be apparent to one of ordinary skill in the art that many changes and modifications can be made thereto without departing from the spirit or scope of the invention as set forth herein.

We claim:

- 1. An isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase enzyme activity and has an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23 or 25-27.
- 5 2. The nucleic acid sequence of claim 1, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23 or 25-27.
 - 3. A vector comprising the nucleic acid sequence of claim 1, wherein the nucleic acid sequence is operably linked to a promoter.
 - 4. A host cell which contains the vector of claim 3.
- 5. The host cell of claim 4, wherein the host cell is selected from the group consisting of a bacterial cell, an algal cell, a yeast cell and a plant cell.
 - 6. The host cell of claim 4, wherein the host cell is a photosynthetic cell.
 - 7. An isolated and/or purified protein having lycopene ϵ -cyclase enzyme activity and having an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23 or 25-27.
 - 8. The protein of claim 7, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23 or 25-27.



AMENDED CLAIMS

[received by the International Bureau on 15 November 1999 (15.11.99); original claims 1,2,7 and 8 amended; remaining claims unchanged (1 page)]

- 1. An isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase enzyme activity and has an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23, 25 or 26.
- 5 2. The nucleic acid sequence of claim 1, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23, 25 or 26.
 - 3. A vector comprising the nucleic acid sequence of claim 1, wherein the nucleic acid sequence is operably linked to a promoter.
 - 4. A host cell which contains the vector of claim 3.
- 5. The host cell of claim 4, wherein the host cell is selected from the group consisting of a bacterial cell, an algal cell, a yeast cell and a plant cell.
 - 6. The host cell of claim 4, wherein the host cell is a photosynthetic cell.
- 7. An isolated and/or purified protein having lycopene ∈-cyclase enzyme activity and having an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23, 25 or 26.
 - 8. The protein of claim 7, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23, 25 or 26.

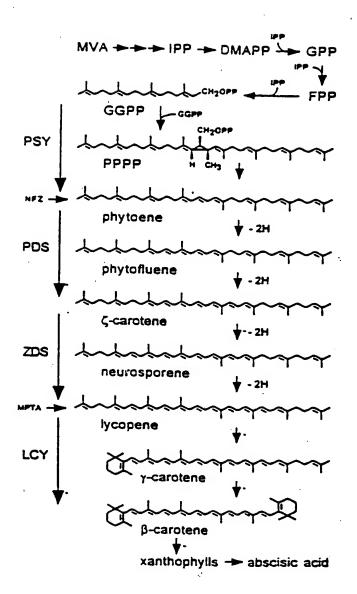
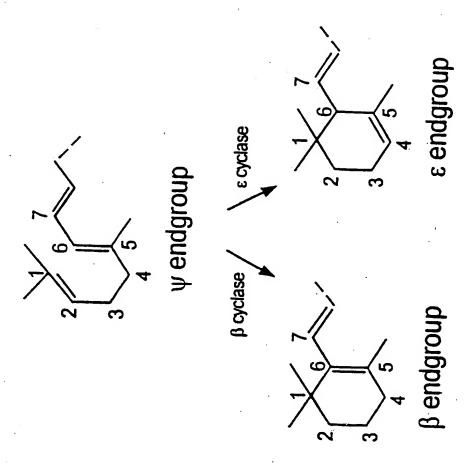


FIGURE 1

£.

FIGURE 2



Arabidopsis thaliana epsilon cyclase:

FIGURE 4

getettete etecteetet accgatttee gaeteegeet eccgaaatee ttatccggat teteteegte tettegattt aaacgetttt etgtetgtta 51 cgtcgtcgaa gaacggagac agaattctcc gattgagaac gatgagagac 101 eggagageae gageteeaca aacgetatag acgetgagta tetggegttg 151 201 cgtttggcgg agaaattgga gaggaagaaa tcggagaggt ccacttatct aatcgctgct atgttgtcga gctttggtat cacttctatg gctgttatgg 251 ctgtttacta cagattetet tggcaaatgg agggaggtga gateteaatg 301 ttggaaatgt ttggtacatt tgctctctct gttggtgctg ctgttggtat 351 ggaattotgg gcaagatggg otcatagage totgtggcac gottototat 401 ggaatatgca tgagtcacat cacaaaccaa gagaaggacc gtttgagcta 451 501 aacgatgttt ttgctatagt gaacgetggt ecagegattg gteteetete ttatggattc ttcaataaag gactcgttcc tggtctctgc tttggcgccg 551 ggttaggcat aacggtgttt ggaatcgcct acatgtttgt ccacgatggt 601 ctcgtgcaca agcgtttccc tgtaggtccc atcgccgacg tcccttacct 651 ccgaaaggte gccgccgctc accagctaca tcacacagac aagttcaatg 701 gtgtaccata tggactgttt cttggaccca aggaattgga agaagttgga 751 ggaaatgaag agttagataa ggagattagt cggagaatca aatcatacaa 801 asaggeeteg ggeteegggt egagttegag ttettgaett tasacsagtt 851 ttaaatccca aattcttttt ttgtcttctg tcattatgat catcttaaga 901 951 cggtct

FIGURE 5

 $\sigma^{i_{\underline{i}}}$

| V | |
|--------|---|
| [1 | , |
| 2 | |
| | |
| C |) |
| 년 대 | • |

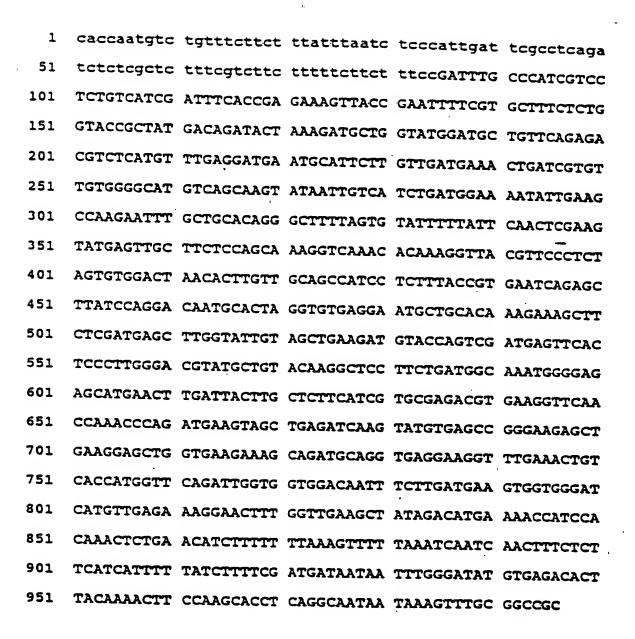
| A. chal. | | 9755 | 6755 SSTDFRLRLP KSLSGFSPSL RFKLSVCTV VERRQHSPI EXDERPESTS STIALDAFAL | Kslsgpspsl | RFRAFSVCTV | VEERROHSPI | EXDEPLETS | TUNGINGUS 419 |
|--|--|--|--|---|--|---|--|---|
| A.thal. Alical. A.aurant. E.berb. E.ured. Consensus | ALELAEIGER | | KKSERSTYLI ANGLESPGIT | SHAWHAWTR | 18 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 | SPCENFOTPA HTOFL HTNFL HGNSL HGNIMWAL | LSVGAAVCHE IVVATVLVHE IVVLSVTAHE IVFLSVTAHE | IUU FYARNAHAL LTAYSVICHI GLAATHOYI VIAALAHEYI |
| | | | Predicted TM helix | helt | | Predi | Predicted TM helix | |
| A. chal. | WHASE, WROTH | SSHHXPREGE | PELNDVPALV | HAGPAIGLLS | YGPPHKGLVP VGAYMMPVLM | GLCFGAGLGI | TVPGLAYMPY | HDGLVHXBLP |
| A. aurant. E. herb. | HHGPLOWGHH HHG. NGNRHH | | LEDONDLYGLV | PAVIATVLFT | VGWINGAPVLA | MI ALGR | TWELLYPLY | HOGLVHORND |
| E. ured. Consensus | HAG. NGNGNA -H1-NH | -SHIPPEGA | FEVNDLYAVV FAALSILLTY (E-NDe-V -Ael-L | PANSILLY -Aal-L | Lastampla | MIGAGN | | HDGLVHQRWP HDGLVH-R-P |
| | | | a a | Predicted TM helix | ا ا.ع | Predicted TM helix | M helix | |
| A. chal. Alical. A. auranc. S. berb. B. ured. Consensus | VGPLADVPTL FRYIPRGYT FRYIPRGYA FHVIPRRGYL FRYIPRKGYL | RXYAAAHQLH RRLYQAHRLH RRLYQAHRLH KRLYVAHRLH KRLYVAHRHH | HTDKTHGV HAVEGEDHCV HAVEGEDHCV HAVRGREGCV HAVRGKEGCV | PYGLFLGPKE SPGFIYAP. SPGFIYAR. SPGFIYAR. | LZEVOGNEEL VDICKQDLOR VDICKQDLOR PADICATIAE | DICEI SRRIKS SGVLRPQDER SGVLRAEAGE RHGRPPICEDA RHGARAGA | YYYASGSGSS PS• RT• AYDREDBASP ANDAGGEDE | 5.5.5 5.5.5 5.5.5 7.5.5 |

7 / 30

ccacgggtcc gcctccccgt ttttttccga tccgatctcc ggtgccgagg 1 actcagctgt ttgttcgcgc tttctcagcc gtcaccatga ccgattctaa 51 101 cgatgctgga atggatgctg ttcagagacg actcatgttt gaagacgaat gcattctcgt tgatgaaaat aatcgtgtgg tgggacatga cactaagtat 151 201 aactgtcatc tgatggaaaa gattgaagct gagaatttac ttcacagagc tttcagtgtg tttttattca actccaagta tgagttgctt ctccagcaac 251 301 ggtcaaaaac aaaggttact ttcccacttg tgtggacaaa cacttgttgc 351 agccatcctc tttaccgtga atccgagctt attgaagaga atgtgcttgg 401 tgtaagaaat gccgcacaaa ggaagctttt cgatgagctc ggtattgtag 451 cagaagatgt accagtcgat gagttcactc ccttgggacg catgctttac 501 aaggcacett etgatgggaa atggggagag cacgaagttg actatetact 551 cttcatcgtg cgggatgtga agcttcaacc aaacccagat gaagtggctg 601 agatcaagta cgtgagcagg gaagagctta aggagctggt gaagaaagca 651 gatgctggcg atgaagctgt gaaactatct ccatggttca gattggtggt 701 ggataatttc ttgatgaagt ggtgggatca tgttgagaaa ggaactatca ctgaagctgc agacatgaaa accattcaca agctctgaac tttccataag 751 ttttggatct tccccttccc ataataaaat taagagatga gacttttatt 801 gattacagac aaaactggca acaaaatcta ttcctaggat tttttttgc 851 tttttattta cttttgattc atctctagtt tagttttcat cttaaaaaaa 901 951 aaaa

÷:





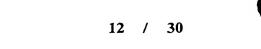


CTCGGTAGCT GGCCACAATC GCTATTTGGA ACCTGGCCCG GCGGCAGTCC GATGCCGCGA TGCTTCGTTC GTTGCTCAGA GGCCTCACGC ATATCCCCCG CGTGAACTCC GCCCAGCAGC CCAGCTGTGC ACACGCGCGA CTCCAGTTTA 101 AGCTCAGGAG CATGCAGATG ACGCTCATGC AGCCCAGCAT CTCAGCCAAT 151 CTGTCGCGCG CCGAGGACCG CACAGACCAC ATGAGGGGTG CAAGCACCTG 201 251 GGCAGGCGGG CAGTCGCAGG ATGAGCTGAT GCTGAAGGAC GAGTGCATCT 301 TGGTGGATGT TGAGGACAAC ATCACAGGCC ATGCCAGCAA GCTGGAGTGT 351 CACAAGTTCC TACCACATCA GCCTGCAGGC CTGCTGCACC GGGCCTTCTC TGTGTTCCTG TTTGACGATC AGGGGCGACT GCTGCTGCAA CAGCGTGCAC 401 GCTCAAAAAT CACCTTCCCA AGTGTGTGGA CGAACACCTG CTGCAGCCAC 451 501 CCTTTACATG GGCAGACCCC AGATGAGGTG GACCAACTAA GCCAGGTGGC CGACGGAACA GTACCTGGCG CAAAGGCTGC TGCCATCCGC AAGTTGGAGC 551 ACGAGCTGGG GATACCAGCG CACCAGCTGC CGGCAAGCGC GTTTCGCTTC 601 CTCACGCGTT TGCACTACTG TGCCGCGGAC GTGCAGCCAG CTGCGACACA 651 ATCAGCGCTC TGGGGCGAGC ACGAAATGGA CTACATCTTG TTCATCCGGG 701 CCAACGTCAC CTTGGCGCCC AACCCTGACG AGGTGGACGA AGTCAGGTAC 751 GTGACGCAAG AGGAGCTGCG GCAGATGATG CAGCCGGACA ACGGGCTGCA 801 ATGGTCGCCG TGGTTTCGCA TCATCGCCGC GCGCTTCCTT GAGCGTTGGT 851 GGGCTGACCT GGACGCGGCC CTAAACACTG ACAAACACGA GGATTGGGGA 901 ACGGTGCATC ACATCAACGA AGCGTGAAAG CAGAAGCTGC AGGATGTGAA 951 GACACGTCAT GGGGTGGAAT TGCGTACTTG GCAGCTTCGT ATCTCCTTTT 1001 TCTGAGACTG AACCTGCAGT CAGGTCCCAC AAGGTCAGGT AAAATGGCTC 1051 GATAAAATGT ACCGTCACTT TTTGTCGCGT ATACTGAACT CCAAGAGGTC 1101 1151 ΑΑΑΑΑ ΑΑΑΑΑΑΑ



| 1 | CTCGGTAGCT | GGCCACAATC | GCTATTTGGA | ACCTGGCCCG | GCGGCAGTCC |
|------------------|------------|-------------|-------------|--------------|------------|
| 51 | GATGCCGCGA | TGCTTCGTTC | GTTGCTCAGA | GGCCTCACGC | ATATCCCGCG |
| 101 | CGTGAACTCC | GCCCAGCAGC | CCAGCTGTGC | ACACGCGCGA | CTCCAGTTTA |
| 151 | AGCTCAGGAG | CATGCAGCTG | CTTTCCGAGG | ACCGCACAGA | CCACATGAGG |
| 201 | GGTGCAAGCA | CCTGGGCAGG | CGGGCAGTCG | CAGGATGAGC | TGATGCTGAA |
| 251 | GGACGAGTGC | ATCTTGGTAG | ATGTTGAGGA | CAACATCACA | GGCCATGCCA |
| 301 | GCAAGCTGGA | GTGTCACAAG | TTCCTACCAC | ATCAGCCTGC | AGGCCTGCTG |
| 351 | CACCGGGCCT | TCTCTGTGTT | CCTGTTTGAC | GATCAGGGGC | GACTGCTGCT |
| 401 | GCAACAGCGT | GCACGCTCAA | AAATCACCTT | CCCAAGTGTG | TGGACGAACA |
| 451 | CCTGCTGCAG | CCACCCTTTA | CATGGGCAGA | CCCCAGATGA | GGTGGACCAA |
| 501 | CTAAGCCAGG | TGGCCGACGG | AACAGTACCT | GGCGCAAAGG | CTGCTGCCAT |
| 551 | CCGCAAGTTG | GAGCACGAGC | TGGGGATACC | AGCGCACCAG | CTGCCGGCAA |
| 601 | GCGCGTTTCG | CTTCCTCACG | CGTTTGCACT | ACTGTGCCGC | GGACGTGCAG |
| 651 | CCAGCTGCGA | CACAATCAGO | GCTCTGGGGC | GAGCACGAAA | TGGACTACAT |
| 701 | CTTGTTCATC | CGGGCCAACG | TCACCTTGGC | GCCCAACCCT | GACGAGGTGG |
| 751 [.] | ACGAAGTCAG | GTACGTGACG | CAAGAGGAGC | TGCGGCAGAT | GATGCAGCCG |
| 801 | GACAACGGGC | TTCAATGGTC | GCCGTGGTTT | CGCATCATCG | CCGCGCGCTT |
| 851 | CCTTGAGCGT | TGGTGGGCTG | ACCTGGACGO | GGCCCTAAAC | ACTGACAAAC |
| 901 | ACGAGGATTO | GGGAACGGT | CATCACATCA | ACGAAGCGTG | AAGGCAGAAG |
| 951 | CTGCAGGATO | TGAAGACAC | TCATGGGGT | GAATTGCGTA | CTTGGCAGCT |
| 1001 | TCGTATCTC | C TTTTTCTGA | ACTGAACCT | CAGAGCTAGA | GTCAATGGTG |
| 1051 | CATCATATT | C ATCGTCTCT | C TTTTGTTTT | A GACTAATCTO | TAGCTAGAGT |
| 1101 | CACTGATGA | A TCCTTTACA | A CTTTCAAAA | AAAAA | |

| HP04 | 1 MLRSLLRGLT | HIPRVNSAQQ | PSCAHARLQF | KLRSMQMTLM | 50 OPSISANTSD |
|----------|--------------------------|--------------------|-----------------------|--------------------------|--------------------------|
| HP05 | MLRSLLRGLT MSVSSLFNLP | HIPKYNSAQQ | PSCAHARLQF | KLRSMQLL | •••••• |
| C brew | MS.SSMLNFT | .LIRLRSLAASRIVSLPL | LSSSFSSFRF | AHRPLSSIS. | PRKLPNFRAF |
| ATOPS | | .TGPPPRFFP | LSSPPSRVHL IRSPVPRTQL | PLCFFSPISL | TORFSAKLTF |
| S ceres. | | PHGAVSSYAK | | FVRAFSAV | • • • • • • • • • |
| 2 (650. | | I HOW TOO IM | PAGUGIPEDI | LEEFPEIIPL | QQRPNTR |
| | 51 | • | | | 100 |
| | AEDRTDHMRG | ASTWAGGQSQ | DELMLKDECI | LVDVEDNITG | 100 HASKLECHKF |
| | SEDRTDHMRG | ASTWAGGQSQ | DELMLKDECI | LVDVEDNITG | HASKLECHKF |
| | SGTA.MTD | TKDAGMDAVQ | | LVDETDRVVG | HVSKYNCHLM |
| | SSQATT.MGE | VVDAGMDAVQ | RRLMFEDECI | LVDENDKVVG | HESKYNCHLM |
| | T.MTD | SNDAGMDAVQ | | LVDENNRVVG | HDTKYNCHLM |
| • | SSETSNDESG | ETCFSGHDEE | QIKLMNENCI | VLDWDDNAIG | AGTKKVCHLM |
| | 101 | | | | • |
| | | PARSUET FOR | OCRETEORN | RSKITFPSVW | 150 |
| | LPHQPAGLLH | RAFSVFLFDD | OCRITIOORA | RSKITFPSVW | |
| | ENIEAKNLLH | | | NTKVTFPLVW | TNTCCSHPLH |
| | EKIESENLLH | RAFSVFLFNS | KYELLLOORS | ATKVTFPLVW | TNTCCSHPLY TNTCCSHPLY |
| | EKIEAENLLH | RAFSVFLFNS | KYELLLQQRS | | TNTCCSHPLY |
| | | | QGELLLOORA | TEKITFPDLW | TNTCCSHPLC |
| | | | | | 2000 20 |
| | 151 | | | | 200 |
| | GQTPDEVDQL | | AKAAAIRKLE | HELGIPAHQL | PA.SAFRFLT |
| | GQTPDEVDQL | SQVADGTVPG | AKAAAIRKLE | HELGIPAHQL | PA.SAFRFLT |
| | RE | SELIQUNALG | VRNAAQRKLL | DELGIVAEDV | PV.DEFTPLG |
| | RE | | VRNAAQRKLL | DELGIPAEDL | PV.DQFIPLS |
| | | KGKLDDKIKG | A LUIN Y LUINT D | DELGIVAEDV HELGIPEDET | PV.DEFTPLG |
| | | | VIIVYAKKID | RELGIFEDET | KTRGKFHFLN |
| • | 201 | | | 1 | - 250 |
| | RLHYCAADVQ | | EHEMDYILFI | RANVTL | |
| | RLHYCAADVQ | | EHEMDYILFI | RANVTL | |
| | RMLY | .KAPSDGKWG | EHELDYLLFI | VRDVKV | |
| | RILY | .KAPSDGKWG | EHELDYLLFI | | DPNPDEVAEV. |
| | RMLY | . KAPSDGKWG | EHEVDYLLFI | VRDVKL | |
| | KIDI | .MAPSNEPWG | EHEIDYILFY | KINAKENLTV | NPNVNEVRDF |
| | 251 | | | | 20.0 |
| | | MMOPDN | GT.OWSDWEDT | IAARFLERWW | 300 |
| | RYVTQEELRO | MMQPDN | GLOWSPWFRI | IAARFLERWW | ADLDAALNTD |
| • | KYVSREELKE | LVKKADAGEE | GLKLSPWFRL | VVDNFLMKWW | DHVEKGTLVE |
| | KYMNRDDLKE | LLRKADAEEE | GVKLSPWFRL | VVDNFLFKWW | DHVEKGSLKD |
| | KYVSREELKE | LVKKADAGDE | AVKLSPWFRL | VVDNFLMKWW | DHVEKGTITE |
| | KWVSPNDLKT | MFADP | SYKFTPWFKI | ICENYLFNWW | EQLDDLSEVE |
| | 301. | | | | |
| | KHEDWGTVHH | TNFA+ | | | |
| | KHEDWGTVHH | | | | |
| | A. IDMKTIHK | | | • | |
| | A.ADMKTIHK | | | | |
| | A.ADMKTIHK | | | · | 4. |
| | NDRQIHR | ML* | * | | • |
| | | | | | |



WO 99/63055

ccaaaaacaa ctcaaatctc ctccgtcgct cttactccgc catgggtgac gactccggca tggatgctgt tcagcgacgt ctcatgtttg acgatgaatg 51 cattttggtg gatgagtgtg acaatgtggt gggacatgat accaaataca 101 attgtcactt gatggagaag attgaaacag gtaaaatgct gcacagagca 151 201 ttcagcgttt ttctattcaa ttcaaaatac gagttacttc ttcagcaacg 251 gtetgeaace aaggtgacat tteetttagt atggaceaac acetgttgea 301 gecatecaet etacagagaa teegagettg tteeegaaac geetgagaga 351 401 451 501 551 601 651 xxxxxxxxx xxxxxxxx xxxxxxxxx tcatgtgcaa aagggtacac 701 tcactgaatg caatttgata tgaaaaccat acacaagctg atatagaaac acacecteaa ecgaaaagea ageetaataa ttegggttgg gtegggteta 751 ccatcaattg ttttttttt ttaacaactt ttaatctcta tttgagcatg 801 ttgattcttg tcttttgtgt gtaagatttt gggtttcgtt tcagttgtaa 851 taatgaacca ttgatggttt gcaatttcaa gttcctatcg acatgtagtg 901 951 atctaaaaaa

FIGURE 12

· , - 1

Plant beta A.c. epsilon Consensus Cyanobacterial enzyme begins -VK--SEALLO LVPETKKEML DFELPMYDD. ...S.Kg-VV DLAVVGGGPA GLAVAQQVSE AGLSVCSIDD VKAGGSEIL. FVOHQQNKDM DEQSKLVDKL PPISIGDGAL DHVVIGCGPA GLAVARSAK LGLKVGLIGD Plant beca A.c. epsilon Consensus VX---S--L- -V------- D--------- D---V-G-GPA GLA-A---- -GL-V--I-P Possible subunit interaction domain Dinucleotide-binding signature 141 Plant beta A.c. epsilon Consensus Conserved region #1 Plant beta VSYLSSKVDS ITEASDGLEL VACDDMMVIP CRLATVASGA ASGELLQYEV GGPRVCVQTA YGVEVEVEMS A.c. epsilon V-----KV-- -----Q-- ----Q-- ----Q-- YG---EV---Consensus Plant beca PFD--KNVfM DWRDsHL-nn -elkeres-i PTVLYAMPYS SNTIVLEETS LVARPGLENd DIQERNVARL PYDPDQMVFM DYRDY..TNE .KVRSLRABY PTVLYAMPHT KSRLFFEETC LAGEDVMPFD LLETKLMLRL A.c. epsilon Consensus ♣ Conserved region #2 ▲ Conserved region #3 Plant beta HLGIKVKAI BEDENCVIPM GGPLPVIPQR VVGIGGTAGA VHPSTGYMVA RTLAAAPVVA NAII-YLGSA DTLGIRILKT YEERMSYIPV GGSLPNTEQK NLAPGAAASN VHPATGYSVV RSLSEAPKVA SVLAEILREE A.c. epsilon --LGI----- -E-E---IP- GG-LP---Q- ----G--A-M VHP-TGY-V- R-L--AP--A --I---L--E Consensus Conserved region #4 Predicted TM helix 421 Plant beta A.c. epsilon Consensus Conserved region #5 Plant beta GGFLGSTLIE GELIVFGLSL FSHASHTSR- EIMIK.GT-P LV-MINNLIQ D-e QGFLGSTLIE GDLVLFALYM FVISPHNIRK GLINHLISDP TGATMIKTYL KV. A.c. epsilon

30

FIGURE 13

-GFL-S-L-- --L--F-L-- F----N--R- -----

Predicted TM helix

Consensus

-



Adonis palaestina E-cyclase cDNA #5 Length: 1898

```
aaaggagtgt totattaatg ttactgtogo attottgcaa cacttatatt
      caaactccat tttcttcttt tctcttcaaa acaacaaact aatgtgagca
   51
      gagtatetgg ctatggaact acttggtgtt cgcaacctca tetettettg
  101
      ccctgtgtgg acttttggaa caagaaacct tagtagttca aaactagctt
  151
  201 ataacataca togatatggt tottottgta gagtagattt toaagtgaga
      gctgatggtg gaagcgggag tagaagttct gttgcttata aagagggttt
      tgtggatgaa gaggatttta tcaaagctgg tggttctgag cttttgtttg
  301
      tocaaatgoa goaaacaaag totatggaga aacaggocaa gotogoogat
  351
     aagttgccac caatacetti tggagaatcc gtgatggact tggttgtaat
  401
  451
      aggittgigga cotgotggte titcactgge tgcagaaget gctaagetag
 501
     ggttgaaagt tggccttatt ggtcctgatc ttccttttac aaataattat
      ggtgtgtggg aagacgagtt caaagatett ggaettgaac gttgtatega
 551
      gcatgcttgg aaggacacca tcgtatatct tgataatgat gctcctgtcc
 601
      ttattggtcg tgcatatgga cgagttagtc gacatttgct acatgaggag
 651
 701·
      ttgctgaaaa ggtgtgtgga gtcaggtgta tcatatctgg attctaaagt
      ggaaaggatc actgaagctg gtgatggcca tagccttgta gtttgtgaaa
 751
     atgagatett tatécettge aggettgeta etgttgeate tggageaget
 801
     tcagggaaac ttttggagta tgaagtaggt ggccctcgtg tttgtgtcca
 851
 901 aaccgcttat ggggtggagg ttgaggtgga gaacaatcca tacgatccca
     acttaatggt attcatggac tacagagact atatgcaaca gaaattacag
     tgctcggaag aagaatatcc aacatttctC tatgtcatgc ccatgtcgcc
1001
     aacaagactt ttttttgagg aaacctgttt ggcctcaaaa gatgccatgc
1051
     cattegatet actgaagaga aaactgatgt cacgattgaa gactetgggt
1101
1151
     atccaagtta caaaagttta tgaagaggaa tggtcatata ttcctgttgg
1201
     tggttettta ccaaacacag agcaaaagaa cctagcattt ggtgctgcag
1251 caagcatggt gcatccagca acaggctatt cggttgtacg gtcactgtca
1301 gaagetecaa aatatgette tgtaattgca aagattttga agcaagataa
1351 ctctgcgtat gtggtttctg gacaaagtag tgcagtaaac atttcaatgc
1401 aagcatggag cagtetttgg ccaaaggage gaaaacgtca aagagcatTe
     tttctttttg gattagaget tattgtgcag ctagatattg aagcaaccag
     aacattettt agaacettet teegettgee aacttggatg tggtggggtt
1501
     teettgggte ticactatea tetitegate tegtetigti ticcatgtae
1551
1601 atgtttgttt tggcgccaaa cagcatgagg atgtcacttg tgagacattt
1651 gctttcagat ccttctggtg cagttatggt aagagettac ctcgaaaggt
1701 agtotoatot attattaaac totagtgttt caccaaataa atgaggatoc
1751 ticquatgig tatatgatca tototatgia tatocigiac totaatotca
     taaagtaaat gccgggtttg atattgttgt gtcaaaccgg ccaatgatat
1851 aaagtaaatt tattgataca aaagtagttt ttttccttaa aaaaaaaa
```

Adonis palaestina ε-cyclase #5 predicted polypeptide TRANSLATE from: 113 to: 1702 Length: 529 amino acids

| 1 51 101 151 201 251 301 351 401 451 | SGSRSSVAYK IPFGESVMDL DEFKDLGLER CVESGVSYLD LEYEVGGPRV EYPTFLYVMP KVYEEEWSYI YASVIAKILK | SSCPVWTFGT EGFVDEEDFI VVIGCGPAGL CIEHAWKDTI SKVERITEAG CVQTAYGVEV MSPTRLFFEE PVGGSLPNTE QDNSAYVVSG | KAGGSELLFV SLAAEAAKLG VYLDNDAPVL DGHSLVVCEN EVENNPYDPN TCLASKDAMP QKNLAFGAAA QSSAVNISMO | QMQQTKSMEK LKVGLIGPDL IGRAYGRVSR EIFIPCRLAT LMVFMDYRDY FDLLKRKLMS SMVHPATGYS AWSSLWPKER | HLLHEELLKR VASGAASGKL MQQKLQCSEE RLKTLGIQVT VVRSLSEAPK KBORAFFI FG |
|---|---|--|--|--|---|
| 451 501 | LELIVQLDIE | SOURT AAR | Q55AVNISMQ RI.PTWMWWGE | AWSSLWPKER | KRQRAFFLFG VLFSMYMFVL |

FIGURE 14



351

DNA sequence of pot 3 cDNA (GenBank R27545) obtained from Nicholas J. Provart potato.seq Length: 1378 August 2, 1996 13:06 Type: N Check: 605 1 tageggnnnn naggatgagt teaaagatet tggtetteaa geetgeattg 51 aacatgtttg gcgggatacc attgtatatc ttgatgatga tgatcctatt 101 cttattggcc gtgcctatgg aagagttagt cgccatttac tgcacgagga 151 gttactcaaa aggtgtgtgg aggcaggtgt tttgtatcta aactcgaaag 201 tggataggat tgttgaggcc acaaatggcc acagtcttgt agagtgcgag 251 ggtgatgttg tgattecetg caggtttgtg actgttgcat cgggagcage 301 ctcggggaaa ttcttgcagt atgagttggg aggtectaga gtttetgtte 351 aaacagetta tggagtggaa gttgaggteg ataacaatee atttgaceeg 401 agcetgatgg ttttcatgga ttatagagae tatgtcagae acgaegetea 451 atetttagaa getaaatate caacatttet etatgecatg eccatgtete 501 caacacgagt ctttttcgag gaaacttgtt tggcttcaaa agatgcaatg 551 ccattcgatc tgttaaagaa aaaattgatg ttacgattga acaccetcgg 601 tgtaagaatt aaagaaattt atgaggagga atggtcttac ataccagttg 651 gaggatcttt gccaaataca gaacaaaaaa cacttgcatt tggtgctgct 701 gctagcatgg ttcatccagc cacaggttat tcagtcgtca gatcactgtc 751 tgaagctcca aaatgcgct tcgtgcttgc aaatatatta cgacaaaatc 801 atagcaagaa tatgcttact agttcaagta ccccgagtat ttcaactcaa 851 gettggaaca etetttggee acaagaacga aaacgacaaa gategttttt 901 cctatttgga ctggctctga tattgcagct ggatattgag gggataaggt 951 cattetecg egegttette egtgtgecaa aatggatgtg geagggattt 1001 ettggtteaa gtetteettn ageagaeete atgttatttg eettetaeat 1051 gtttattatt gcaccaaatg acatgagaag aggettaate agacatett 1101 tatctgatcc tactggtgca acattgataa gaacttatct tacattttag 1151 agtaaattcc tcctacaata gttgttgaan nagaggcctc attacttcag 1201 attcataaca gaaatcgcgg tctctcgagg ccttgtatat aacattttca 1251 ctaggttaat attgcttgaa taagttgcac agtttcagtt tttgtatctg

1351 gtatatataa attttataaa aaaaaaaa TRANSLATE from: 14 to: 1147 poteps.pep Length: 378 1 DEFKDLGLQA CIEHVWRDTI VYLDDDDPIL IGRAYGRVSR HLLHEELLKR 51 CVEAGVLYLN SKVDRIVEAT NGHSLVECEG DVVIPCRFVT VASGAASGKF 51 LQYELGGPRV SVQTAYGVEV EVDNNPFDPS LMVFMDYRDY VRHDAQSLEA 101 151 KYPTFLYAMP MSPTRVFFEE TCLASKDAMP FDLLKKKIML RINTLGVRIK 201 EIYEEEWSYI PVGGSLPNTE QKTLAFGAAA SMVHPATGYS VVRSLSEAPK CAFVLANILR QNHSKNMLTS SSTPSISTQA WNTLWPQERK RQRSFFLFGL 251 ALILQLDIEG IRSFFRAFFR VPKWMWQGFL GSSLSXADLM LFAFYMFIIA 301 PNDMRRGLIR HLLSDPTGAT LIRTYLTF*

1301 cttcttttt gtccaagatc atgtattgan ccaatttata tacattgcca

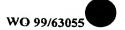
FIGURE 15A

Chimeric lettuce/potato lycopene ε-cyclase: converts lycopene to δ-carotene, the lettuce cDNA converts lycopene to t-carotene and the potato cDNA does not produce an active enzyme

(amino acids in lower case are from lettuce and those in uppercase are from the potato cDNA; an AvaII site in common to the two cDNAs was used to construct the chimera)

```
l mecfgarnmt atmavftcpr ftdcnirhkf sllkqrrftn lsassslrqi
51 kcsaksdrcv vdkqgisvad eedyvkaggs elffvqmqrt ksmesqskls
101 eklaqipign cildlvvigc gpaglalaae saklginvgl igpdlpftnn
151 ygvwqdefig lglegciehs wkdtlwyldd adpirigray grwhrdllhe
201 ellrrcvesg vsylsskver iteapngysl iecegnitip crlatvasga
    asgkfleyel gGPRVSVQTA YGVEVEVDNN PFDPSLMVFM DYRDYVRHDA
251
301 OSLEAKYPTF LYAMPMSPTR VFFEETCLAS KDAMPFOLLK KKLMLRLNTL
351 GVRIKEIYEE EWSYIPVGGS LPNTEQKTLA FGAAASMVHP ATGYSVVRSL
     SEAPKCAFVL ANILRONHSK NMLTSSSTPS ISTOAWNTLW POERKRORSF
401
     FLEGLALILQ LDIEGIRSFF RAFFRVPKWM WQGFLGSSLS XADLMLFAFY
451
    MFIIAPNOMR RGLIRHLLSD PTGATLIRTY LTF*
```

FIGURE 15B



| GAP comparison of Arabidopsis e-cyclase x potato e-cyclase (partial) |
|--|
| Length Weight: 12 Average Match: 2.912 Quality: 1485 Length: 529 Ratio: 3.929 Gaps: 1 |
| Match display thresholds for the alignment(s): = IDENTITY : = 2 . = 1 |
| 151 EDEFNDLGLOKCIEHVWRETIVYLDDDKPITIGRAYGRVSRRLLHEELLR 200 |
| 201 RCVESGVSYLSSKVDSITEASDGLRLVACDDNNVIPCRLATVASGAASGK 250 |
| 251 LLQYEVGGPRVCVQTAYGVEVEVENSPYDPDQMVFMDYRDYTNEKVRSLE 300 |
| 301 AEYPTFLYAMPMTKSRLFFEETCLASKDVMPFDLLKTKLMLRLDTLGIRI 350 |
| 351 LKTYEEEWSYIPVGGSLPNTEQKNLAFGAAASMVHPATGYSVVRSLSEAP 400 . |
| 401 KYASVIAEILREETTKQINSNISRQAWDTLWPPERKRQRAFFLFG 445 |
| 446 LALIVQFDTEGIRSFFRTFFRLPKWMWQGFLGSTLTSGDLVLFALYMFVI 495 |
| 496 SPNNLRKGLINHLISDPTGATMIKTYLKV 524 . .: : : : : 350 APNDMRRGLIRHLISDPTGATMIND |

FIGURE 16





| Adonis | pal | ae | s t | in | a | Ιp | il |
|--------|-----|----|-----|-----|----|----|----|
| | 1 | at | tc | at: | ct | tc | ac |
| | • | - | _ | | | | • |

| 1 | attcatcttc | agcagcgctg | tcgtactctt | tctatatctt | cttccatcac |
|------|------------|------------|-------------|------------|------------|
| 51 | taacagtagt | cgccgacggt | tgaatcggct | attcgcctca | |
| 101 | tgggtgaagt | cactgatgct | ggaatggatg | ctgttcagaa | gcggctcatg |
| 151 | | aatgtatttt | | aatgacaagg | |
| 201 | | tacaactgtc | | aaagatagag | |
| 251 | tocttcacao | agccttcagt | attttcttat | tcaactcaaa | |
| 301 | cttcttcage | aacgateege | cacaaacota | | atatgaattg |
| 351 | aaacacatgt | tacaatcatc | ctctctttc | acattcccgc | tcgtatggac |
| 401 | assattatet | tgcagtcatc | 220000000 | tgattccgag | |
| 451 | Ctaccatta | cggtgtacga | aacgccgcac | aaagaaagct | tttagacgag |
| 501 | teaggeacte | cagctgaaga | tgtcccagtt | gatgaattta | ctcctcttgg |
| 551 | togcattett | tacaaagctc | catctgacgg. | caaatgggga | gagcacgaat |
| | tggactatct | cctatttatt | gtccgagatg | tgaaatacga | tecaaaccca |
| 601 | gatgaagttg | ctgatgctaa | gtatgttaat | cgcgaggagt | tgagagagat |
| 651 | actgagaaaa | gctgatgctg | gtgaagaggg | actcaagttg | tctccttggt |
| 701 | ttagattggt | tgttgataac | tttttgttca | agtggtggga | tcatgtagag |
| 751 | cagggtacga | ttaaggaagt | tgctgacatg | aaaactatcc | acaagttgac |
| -801 | ttaagaggac | ttctctcctc | tgttctacta | tttattttt | gctacaataa |
| 851 | gtgggtggtg | ataagcagtt | tttctgtttt | Ctttaattta | tggcttttga |
| 901 | atttgcctcg | atottoaact | tgtaacatat | ttacacaaat | atgagaga |
| 951 | | tttgaggctg | aatttatatt | tttqqqaaca | tastastas |
| 1001 | 22 | | | cccgggaaca | caacaatgtt |

FIGURE 17A

Adonis palaestina Ipi2

| pu | reescina ip | 12 | | | 1 |
|--------|-------------|------------|------------|------------|------------|
| 1 | ttttaaagct | ctttcgctcc | accaccatca | 220002000 | aatttctctg |
| 51 | tacaaaagtt | aaaaacacco | Ctttgggctt | tagecageca | atatcggaat |
| 101 | ccttgtttac | gatacgcatc | taaaccagta | atteteget | ttaatttgtt |
| 151 | tcctaaatta | ggccctttc | Cogatece | accettgget | cgtcgatcag |
| 201 | gattaatcct | ttatatagra | tetteteese | agaattatgt | acattatcag |
| 251 | cttcqtqttc | tteteceart | gttcatcttc | Caccaccaaa | acattatcag |
| 301 | tctatttctt | cttccatcac | taacactcct | agcagcgttg | regtactett |
| 351 | gttcgcctca | acotcoacta | taacagtcct | cgccgagggt | tgaatcggct |
| 401 | CCGtccagaa | acaacttata | tgggtgaagt | egetgatget | ggtatggatg |
| 451 | aatgacaagg | tratragaca | ttcgacgatg | aatgtatttt | ggtggatgag |
| 501 | aaagatagag | CCaCaaaaac | tgattccaaa | tacaactgtc | atttgatgga |
| 551 | tcaactcaaa | atacaaatta | tgcttcacag | agccttcagt | gttttcttat |
| 601 | acattcccc | testates | cttcttcagc | aacgatctgc | aacgaaggta |
| 651 | tgattccga | ctgtatggae | aaacacctgt | tgcagccatc | ccctcttccg |
| 701 | aaaccaaacc | tttatagaag | aaaattttct | cggggtacga | aacgctgcac |
| 751 | addagaagcc | LLLagacgag | ctaddcattc | cadctdaada | CGTaccages |
| 801 | gargaarrea | crecrettgg | tegeattett | tacaaaactc | Catctgacgg |
| 851 | uuaacyyyya | gagcacgaac | tggactatct | tetattatt | gtccgagatg |
| 901 | cyaaacacya | cccaaaccca | gatgaagttg | ctgacgctaa | atacattaat |
| 951 | cycgaggagt | tgaaagagat | actgagaaaa | gctgatgcag | GEGERATO |
| 1001 | aacaaagttg | ceteettggt | ttagattggt | totogataac | ***** |
| 1051 | ageggegga | tcatgtagag | gaggggaaga | ttaaggacgt | CCCCCCCC |
| | addactattt | acaagttgac | ttaagagaaa | atetettaan | ttctactatt |
| 1101 | caarrrage | ttcaataagt | ggatggtgat | gagcagtttt | tatocttcct |
| 1151 | ccaaccccgg | ctttcaatt | tgctttatgt | gttgaacttg | taacatattt |
| 1201 | agicaaalat | gagaccttgt | gagttgaatt | tgaggttata | tttatagtrr |
| 1251 | EGGGGAACAFA | | | _ ,, | |

FIGURE 17B





÷ ':

Haematococcus pluvialis Ipil

| 1 51 101 151 201 251 301 351 401 451 501 651 701 751 801 | cctttacatg cgacggaaca acgagctggg ctcacgcgtt atcagcgctc ccaacgtcac gtgacgcaag | tgettegtte geccageage catgeagatg cegaggaeage taceacatea tttgaegate cacettecea ggcagaecec gtacetggeg gataceageg tggggegage cttggegeee aggagetgeg | ccagetgtge acgetcatge cacagaceae atgagetgat atcacagge gcetgcagge aggggegaet agtgtgtgga agatgaggtg caaaggetge caccagetge tgccgeggae acgaaatgga aaccetgaeg gcagatgatg | acacgcgcga agcccagcat atgaggggtg gctgaaggac atgccagcaa ctgctgcaac gctgctgcaa cgacaacctaa tgccatccgc cggcaagcgc gtgcagccag ctacatcttg aggtggacga cagccggaca | caagcacetg gagtgcatet getggagtgt gggcettete cagegtgeae etgeageeae gecaggtgge aagttggage gtttegette etgegaeaea tteateeggg agteaggtae |
|--|--|---|---|---|--|
| 701 751 | atcagegete ccaacgteae gtgacgeaag atggtegeeg gggetgacet acggtgeate gacacgteat tetgagaetg | tggggcgagc cttggcgcc aggagctgcg tggtttcgca ggacgcggcc acatcaacga ggggtggaat aacctgcagt accgtcactt | acgaaatgga aaccctgacg gcagatgatg tcatcgccgc ctaaacactg agcgtgaaag tgcgtacttg caggtcccac | gtgcagccag ctacatcttg aggtggacga cagccggaca gcgcttcctt acaaacacga cagaagctgc | ctgcgacaca ttcatccggg agtcaggtac acgggctgca gagcgttggt |

FIGURE 18A

Haematococcus pluvialis Ipi2 tggaacctgg cccggcggca gtccgatgcc gcgatgcttc gttcgttgct cagaggeete acgeatatee egegegtgaa eteegeecag cageccaget 51 gtgcacacge gegaetecag titaagetea ggagcatgca getgettgee 101 gaggaccgca cagaccacat gaggggtgca agcacctggg caggcgggca gtcgcaggat gagctgatgc tgaaggacga gtgcatctta gtggatgctg 151 201 acgacaacat cacaggecat gecageage tggagtgeca caaatteeta 251 ccacatcage ctgcaggect getgcacegg geettetetg tgttcctgtt 301 tgacgaccag gggcgactgc tgctgcaaca gcgtgcacgc tcaaaaatca 351 cetteceaag tgtgtggacg aacacetget geagecacec tetacatggg 401 cagaccccag atgaggtgga ccaactaagc caggtggccg acggcacagt 451 acctggcgca aaagctgctg ccatccgcaa gttggagcac gagctgggga 501 taccagegea ccagetgeeg. gcaagegegt tregetreet caegegittg 551 cactactgtg ccgcggacgt gcagccggct gcgacacaat cagcgctctg 601 gggcgagcac gagatggact acatettatt cateegggee aaegteacet 651 tggcgcccaa ccctgacgag gtggacgaag tcaggtacgt gacgcaagag 701 gagetgegge agatgatgea geeggacaae gggttgeaat ggtegeegtg 751 gtttegeate ategeegege getteettga gegttggtgg getgaeetgg 801 acgeggeeet aaacactgae aaacacgagg attggggaac ggtgeateac 851 901 atcaacgaag cgtgaaggca gaagctgcag gatgtgaaga cacgtcatgg 951 ggtggaattg cgtacttggc agcttcgtat ctcctttttc tgagactgaa cctgcagagc tagagtcaat ggtgcatcat attcatcgtc tctcttttgt 1001 1051 tttagactaa tctgtagcta gagtcactga tgaatccttt acaactttca 1101 aaaaaaaa

FIGURE 18B



BNSDOCID: -WO

99630554115

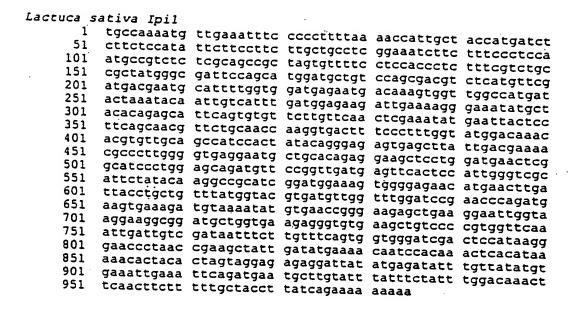


FIGURE 19A

| Lactuca s | ativa Ipi2 | | | | |
|-----------|------------|------------|------------|------------|---|
| 1 | tattcgcttc | aaaatctctt | ccattaactg | ctcaaatctc | caccttcgcc |
| 51 | ggtcttaatc | tccgccggcg | cactttcacc | accataaccg | ccgccatggg |
| 101 | tgacgattcc | ggcatggacg | ctgtccagag | acgtctcatg | tttgatgatg |
| 151 | aatgcatttt | ggttgatgaa | aatgacaatg | ttcttgggca | tgataccaaa |
| 201 | tacaattgtc | acttgatgga | gaagattgag | aaagataatt | tgcttcatag |
| 251 | agcattcagt | gtattttat | tcaattcaaa | atacgaatta | ctccttcagc |
| 301 | aaaggtcaga | aaccaaggtg | acatttcctt | tggtatggac | aaacacctgt |
| 351 | tgcagccatc | cactatacag | agaatcggag | ttaattcccg | aaaatgccct |
| 401 | tggggtcaga | aatgctgcac | agaggaagct | tctagatgaa | ctcggtatcc |
| 451 | ctgctgaaga | tgttccagtt | gatgagttca | caactttagg | togcatgttg |
| 501 | tacaaggctc | catctgatgg | aaaatggggt | gaacatgaag | ttgattacct |
| 551 | actcttcctc | gtgcgtgacg | ttgccgtgaa | cccaaaccct | gatgaggtgg |
| 601 | cggacattag | atacgtgaac | caagaagagt | taaaagagtt | actaaggaag |
| 651 | gcggatgcgg | gtgaggaggg | tttgaaattg | tccccatggt | ttaggctagt |
| 701 | ggtggacaac | ttcttgttca | aatggtggga | tcatgtccaa | aaggggacac |
| 751 | tcaatgaagc | aattgacatg | aaaaccattc | ataagttgat | atgaaaaatg |
| 801 | | atggtggtgg | | | tgttcaagtc |
| 851 | | ttttttaac | | | |
| 901 | | ttgtaacgta | | | |
| 951 | | cgttaattta | | | • |
| | | | | | |

FIGURE 19B

WO 99/63055

Chlamydomonas reinhardtii Ipil

(Note: the isomerase cDNA probably ends at ca. base 1103; the second half of the cDNA is similar to extensin and other hydroxyproline-rich structural proteins)

| 1 51 | ggcacgagc | | t ttaccatga | c atcgggaat | + +aa |
|---------|------------|--------------|-------------|--------------|-------------|
| | aactacctca | attactcaa | taactcgcg | | |
| 101 | cgctgttttc | : tctgctccad | ctaccgage | | |
| 151 | gatgtcataa | actcccacti | atatgagat | | , , , , , , |
| 201 | ccagagcgca | acctgtctta | | | |
| 251 | caaagccgtg | ctctcgttg | | | gcgcctcgcg |
| 301 | aggactttca | caggeteaaa | | | ggccgggagc |
| 351 | cctgggaagg | cacgggcctg | | | agttcgtcaa |
| 401 | tgcttggtgg | | | | gegggaegag |
| 451 | cgactgccac | | | | CCaacaagta |
| 501 | accgcgcctt | | | f ccagccctgc | ggccgcctac |
| 551 | cagcagcgcg | | | ccgacggccg | actgctgctg |
| 601 | ctgctgctcg | | | CCgggtatat | ggaccaacac |
| 651 | Cadeacacac | | | geeggaegag | gtggacctgc |
| 701 | cggcggcggt | | | gcatcaaggg | ggcggcggtg |
| 751 | cgcaagctgc | | ggggataccg | ccggagcagg | ttcccgcctc |
| 801 | ctccttctcc | 3 - | gtctgcacta | ctgcgccgcc | |
| 851 | cgcacggccc | ggcggcggag | tggggcgagc | acgaggtgga | ,,,,, |
| 901 | ttcgtgcggc | cgcagcagcc | cgtcagcctg | cagcccaacc | ctacgtgctg |
| 951 | ggacgccacg | cgctacgtga | cgctgccgga | gcttcagtcc | cagacgaggt |
| 1001 | accccggcct | cagctggagc | ccctggttcc | gcatcctggc | atgatggcgg |
| | gccttcctgc | ccgcctggtg | gggcgacctg | | cacacagccc |
| 1051 | cggcagccga | Ctgtcggact | ggggcaccat | | ggcgcccggg |
| 1101 | aaaggggaag | caggggcggg | agcgggggat | ccaccgcgtc | atgtgaagaa |
| 1151 | ttgtgatgcg | gcgtgggatg | aggtctgaag | gaatgggaat | gtgaatgcga |
| 1201 | cgggcgtgag | cgtgtgtgta | cgtgagcgac | acagggggaa | aatcgggggg |
| 1251 | gcgatgggta | catgtgtgtg | cggagggtcg | aaagccggga | ggcggaccgc |
| 1301 | gcatagcgtg | ttgtgtgtgt | geggetgege | gtgggtcggt | cggttgcgcg |
| 1351 | acggaggaga | aggcacacgc | aggtggcgcg | gggtatgtgg | gcacccgggc |
| 1401 | ggcgggcctc | actcctggtc | atacceases | gaggtgtgtc | aggggccatg |
| 1451 | ggggctgcac | ccatatgage | gtgcccagtg | gtctcgtggg | cagagtggca |
| 1501 | tcacttggtg | aggtggggcg | ggcgcactgc | cgcgctgggc | taagtcctta |
| 1551 | gaaggacacg | gtgtgtgagc | aggtggctgt | gggcggcggg | cgcagtggca |
| 1601 | ggcggatagc | gatatgacgt | ggtggagctc | tggccgtgcc | ggccgtgagg |
| 1651 | tgcaggccgc | | tgtgcttggc | cgctgtaatg | cgggagaatg |
| 1701 | cgttggggag | gagaagcggg | cggtggcagg | aggccgcagg | ctgcagcacc |
| 1751 | gggcgcctga | gtgccgcctg | caggegegge | gccgggcggg | cctgagtaat |
| 1801 | ggacgagetg | gtagtggcgg | ccacaggagg | cgcaggaggc | agcagcagga |
| 1851 | gtggccatac | gagggacccg | ttggcaaccc | | gtgtaacata |
| | 3-39-catac | aaaaaaaa | aaaa | _ | |

FIGURE 20



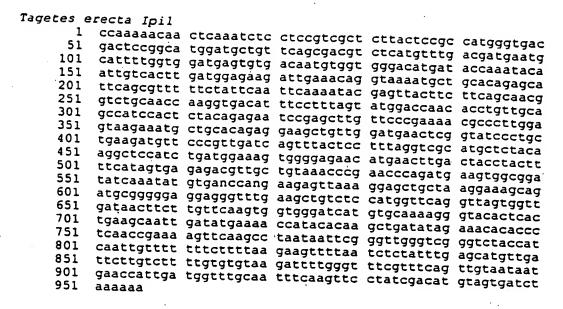


FIGURE 21A

| Oryza | sat | iva Ipil | | | | |
|-------|---|---|---|---|--|---|
| | 1 51 101 151 201 201 301 301 401 451 501 551 601 701 751 801 | agcggctcat gttgttggc atctgaaaat aatatgaact ctagtttgga gcttatacag tcttggatga acccctcttg tgaacacgag tcccgaaccc ctgaaggagc gtctccctgg atcacgtcga cacaagctga aagactctgt gaagtcagaa | gttcgacgac atgaatcaaa ctacttcata cctactccag ccaacacttg gaaaactacc gctgggcatc gtcggatgct cttgactacc ggacgaagtg tcatccgcaa ttccggctgg gaaaggcacc agtaaggact tcttgtctg gaagcttttg | gaatgcattt atataactgc gggctttcag caacgatctg ctgcagcat ttggtgttag ccagctgaag ttacaaggcc tgctgttcat gccgatgtga agcggacgcc ttgttgacaa ctcaacgagg gcgatgtgt ctgcatatta tatgtttctg | cgggatggac tggtggatga catctgatgg tgtattcctg caacaaaggt cctctgtacc aaatgctgct atgtgccagt ccatctgatg cgtccgcgac aatacgtgag ggagaggaag cttcctcatg ccgtggacat ggctggacat ggctggacag | acaagacaat aaaaaatcga ttcaactcaa tacatttcct gtgagtctga cagaggaagc tgaccaattc gaaaatgggg gtgaaggtag ccgtgaagcag gcctgaagct ggctggtggg ggagaccatc aatgatcctg ggaagttgca |
| 9 | 901 951 901 | gcaaacttct | tgactgagag atattataca | attcccttat | agagtgtcta | tgttaattta |

FIGURE 21B

<u>1.</u>

| Isomerases (IPI) or Complemtation in E.coli | 61 -HGDDSGHDAVQR RLHFDDECILVDECD AHGDDSGHDAVQR RLHFDDECILVDEND THGDDSGHDAVQR RLHFDDECILVDEND THGEVADAGHDAVQK RLHFDDECILVDEND HGEVTDAGHDAVQK RLHFDDECILVDEND AAAAVEDAGHDAVQK RLHFDDECILVDEND AHTDTROAGHDAVQK RLHFDDECILVDEND AHTDTROAGHDAVQR RLHFDDECILVDEQD HTDSAUBAGHDAVQR RLHFDDECILVDEDH HRGASTWAG-GQSQD ELHLKDECILVDVEDH HRGASTWAG-GQSQD ELHLKDECILVDVEDH HRGASTWAG-GQSQD ELHLKDECILVDVEDH ARGASTWAG-GQSQD ELHLKDECILVDVEDH HRGASTWAG-GQSQD ELHLKDECILVDVEDH HRGASTWAG-GQSQD ELHLKDECILVDVEDH HASSSTHEGTGLSQD PRAVORECILVDVEDH | | 241 255 256 270 VAUNPUBEUADIKY VSHEEIKELLERADA VGLOPHPDEUADIKY VSHEEIKELLERADA VANNPHPDEUADIKY VNDEEIKELIRRADA VKYDPHPDEUADIKY VNREELKEILRRADA VKYDPHPDEUADKY VNREELKEILRRADA VKVOPHPDEUADKY VSREELKEILRRADA VKVOPHPDEUADKY VSREELKEILRRADA VKLQPHPDEUARKY VSREELKEILRRADA VKLQPHPDEUARKY VSREELKEILRRADA VTLAPHPDEUDEUKY VTQEELRQHAQP VSLQPHPDEVDEVRY YTGEELRQHAQP VSLQPHPDEVDEVRY VTQEELRQHAQP | Tagetes erecta (marigold) Lactuca sativa (romaine lettuce) Lactuca sativa (romaine lattuce) Adonis palaestina (pheasant's eye) Oryza sativa (rice). Arabidopsis thaliana Haematococcus pluvialis Haematococcus pluvialis Chlamydomonas reinhardtii |
|--|---|---|--|---|
| ClustaiM 1.7 Multiple Sequence Alignment of Plant and Green Algal Isopentenyl Pyrophosphate Isomerases (IPI) These amino acid sequences were predicted by CDNAs that were isolated and identified by color Complemtation in | RKSSFPPHPSLAAASVELHPLSSA DRCRTLS1SSSIT NSPRRGLNRLFASTS IRPLSSISPRKLPNFRAFSGT KLRSHQLLS | VELENSKYELLLOOR SATKUTFPLUMTHYC CSHP VELENSKYELLLOOR SATKUTFPLUMTHYC CSHP VELENDOGGRILLOOR SATKUTFPLUMTHYC CSHP VELENDOGGRILLOOR ARSKITFPSUMTHYC CSHP VELENDOGGRILLOOR ARSKITFPSUMTHYC CSHP VELENDOGGRILLOOR ARSKITFPSUMTHYC CSHP VELENDOGGRILLOOR ARSKITFPSUMTHYC CSHP | 226 EHELDYLLFIVRD EHELDYLLFIVRD EHELDYLLFIVRD EHELDYLLFIVRD EHELDYLLFIVRD EHELDYLLFIVRD EHELDYLLFIVRD EHELDYLLFIVRD EHEVDYLLFIVRD EHEWDYLLFIVRD EHEWDYLLFIRRN EHEWDYLLFIRRN | 301 315 316 GTLTEAIDNKTI HKLI 232 Tagec GTLNEAIDNKTI HKLT 280 Lactu GTLNEAIDNKTI HKTT 293 Adoni GTI |
| Sequence Alignment of Plant and Gr inces were predicted by cDNAs that | 15 16 30 31 HSSIRINPLYSIFST TTKTLSASCSSPAVH LQ | 105 106 120 120 VGHOTKYNCHME KIETGKHLHANFS VGHOTKYNCHME KIEKGMHLHANFS VGTOSKYNCHME KIEKDHLHANFS VGTOSKYNCHME KIEKDHLHANFS VGHOSKYNCHME KIEKHLLHANFS VGHOSKYNCHME KIEKHLLHANFS VGHOSKYNCHME KIEKHLLHANFS TGHASKLECHKEL PHQPAGLLHANFS TGHASKLECHKEL PHQPAGLLHANFS TGHANKYDCHRE AAKGQPGGRLHANFS | 196 VPUDEFPLGRALY- VPUDEFFPLGRILY- VPUDEFFPLGRILY- VPUDEFFPLGRILY- VPUDEFFPLGRILY- VPUDEFFPLGRILY- VPUDEFFPLGRILY- VPUDEFFPLGRILY- VPUDEFFPLGRILY- LPASARRELRHINC LPASARRELRHINC VPASSFSFLTRLHYC | 271 GEGUKLSPWFRLVV DNFLFKWUDHVOK GTL GEGUKLSPWFRLVV DNFLHKWUDHVOK GTL GDEGUKLSPWFRLVV DNFLHKWUDHVOK GTL GDEGUKLSPWFRLVV DNFLHKWUDHVOK GTL GDEGUKSPWFRIJA ARFLFRWADLDA ALH -DNGLOWSPWFRIJA ARFLFRWADLDA ALH -DPGLSWSPWFRIJA TQPAFLPAWGDLKR RWR |
| Clustaiw 1.7 Multiple These amino acid seque | 1 T.erocra 1 2 L.saciva 1 3 L.saciva 2 4 A.polaestina 2 5 A.palaestina 1 6 O.saciva 1 7 A.thaliana 1 8 A.thaliana 2 9 H.pluvialis 1 10 H.pluvialis 2 11 C.reinhardtil 1 | 1 T.erecta 1 2 L.sativa 1 3 L.sativa 2 4 A.palaestina 2 5 A.palaestina 1 6 O.sativa 1 7 A.thallana 1 8 A.thallana 2 9 H.pluvialis 1 10 H.pluvialis 2 11 C.reinhardtil 1 | 1 T.erecta 1 2 L.sativa 1 3 L.sativa 2 4 A Palaestina 2 5 A-palaestina 2 6 O.sativa 1 7 A thaliana 1 9 A.thaliana 1 10 H.pluvialis 1 11 C.reinhardtii 1 | 1 Terecta 1 2 L.sativa 1 3 L.sativa 1 5 L.sativa 1 5 G 6 6 O.sativa 1 5 G |

```
Comparison using GAP program of the Genetics Computer Group
         Gap Weight:
                              50
                                        Average Match: 10.000
        Length Weight:
                                       Average Mismatch:
                                                                0.000
               Quality:
                             17392
                                                     Length:
                                                                  1904
                  Ratio:
                             9.411
                                                       Gaps:
Percent Similarity: 95.331
                                      Percent Identity: 95.331
Match display thresholds for the alignment(s): | = IDENTITY
 Adonis palaestina E-cyclase #3 x Adonis palaestina E-cyclase #5
        1 gagagaaaaigagtgttatittaatgttactgtcgcattcttgcaacac: 49
1 .....aaaggagtgttctattaatgttactgtcgcattcttgcaacact 44
       99 tga.cggagtatctagctatggaactacttggtgttcgcaacctcat
95 tgagcagagtatctggctatggaactacttggtgttcgcaacctcat
           ttcttgccctgtctggacttttggaacaagaaaccttagtagttcaaaac 197
      148
      145
     198 tagettataácatacategátatggttettettgtagagtagatttteaa 247
195 tagettataácatacategatatggttettettgtagagtagatttteaa 244
     248 gtgagggctgatggtggaagcgggagtagaagttctgttgcttataaaga
245 gtgagagctgatggtggaagcgggagtagaagttctgttgcttataaaga
     298 gggttttgtggatgaagagattttatcaaagctggtggttctgagcttt
          tgtttgtccaaatgcagcaaacaaagtctatggagaaacaggccaagctc
     348
     345
          gccgataagttgccaccaatacettttggagaatctgtgatggacttggt 447
     398
     395
          tgtaataggttgtggacctgctggtctttcactggctgcagaagctgcta 497
     448
     445
     498 agctagggttgaaagttggccttattggtcctgatcttcctttta
     548 aattatggtgtgtgggaagacgagttcaaagatcttggacttgaacgtt
545 aattatggtgtgtgggaagacgagttcaaagatcttggacttgaacgtt
          tatcqaqcatqcttqqaaqqacaccatcqtatatcttqacaatqatqctctatccgaqcatqcttggaaqqacaccatcqtatatcttgataatqatqctc
     598
     595
    648 ctgtccttattggtcgtgcatatggacgagttagtcgacatttgctacat 697
    698 qaaqaqttqctqaaaaqqtqtqtcqaqtcaqqtqtatcatatctqaattc
          taaaqtqqaaaqqatcactqaaqctqqtqatqqccataqccttqtaqt
    798 grgaaaacgacarcrrratcccrrgcaggertgcractgractgratcrgga 847
795 grgaaaargagatctttatcccrrgcaggertgcractgractgratcrgga 844
```

FIGURE 23

WO 99/63055

FIGURE 23 (cont.)

| 84 | T 1 1 | 1951 | FFF | ago | ggą | ąą | ÇŢ | ÇÇ | ţg. | ga | gta | ato | aa | σt | aσ | at. | aaa | ٠÷، | ٠ + ~ | ~= | | ٠. | |
|--------------|--------------|------------|-----------------|--------|---------------------------------------|---------------------------------------|--|---------------------------------------|--------|---------------------------------------|--------|---|--|--|--|----------|---------------------------------------|--|----------------|--|---|------------|-----------|
| 84. | 5 ģċá | gċ | ξęċ | aģ | ģģå | aa | cti | FF | få | gå | gta | f | aa | ΪĮ | Ιľ | ĮĮ. | ֓֓֓֓֓֓֓֟֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓ | וַוַן | ŢŢ | 71 | 1111 | • | 397 |
| 898 | 8 F9F | çç | aa | çŧç | 765 | ţą | Fgg | g t c | jţ | gga | agg | , t | gá | aa | ta | σa | יככ זמנ | iċa | a + | 90 | 9 C C C | | 94 |
| 89 | _ | ĊĊ | ààà | င်ငင့် | jċt | tå | Łģ | ggg | 340 | կ գե | ago | ιŁŁ | ga | ďα | ff | ďΑ | TI I | | ŢĮ | II. | ֓֞֞֞֞֓֓֞֞֓֓֞֜֞֜֓֓֡֓֓֓֓֡֓֡֓֡֓֡֓֡֡֡֡֓֓֡֡֡֡֡֡֓֓֓֡֡֡֡֡֡֡֡ | Γ. | 47 |
| 948 | ato | FF | aç | FĖŦ | af | gg: | Fat | fţi | tai | ¢ g | jąç | ţa | çå | gą | ga | cta | ata | ιĖα | ca | ac | 2022 | y : | 944 |
| 945 | atc | ĊĊā | iáċ | tta | aat | ģģ | ţáł | F | cai | ξģ | Jac | :Ła | ca. | ga | ga | ξť | | ίſ | IJ | ֓֞֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓ | atac agaa agaa | | 94 |
| 998 | tta | cac | 159 | ç † ç | gg | aad | ga q | g | a | Fat | -66 | ąą | çå | ţţ | - FÇI | ţ¢ţ | at | ġŧ | ca | tac | gaa. CCa CCa | 2 3 F 1 | 047 |
| 995 | tta | Ċáġ | rtġ | ċŧċ | ġġ | ààq | jáá | ıġå | ia | Lat | çç | ala | ca: | ff. | fç | Łċ | at | qf | ĮΤ. | | | ן ז | 044 |
| 1048 | 959 | 966 | aac | ráa | ga | 951 | 777 | F | 7 | 99 | igg | ąą | ąċ | çţ | gţţ | FFG | gç | ċt | caa | aaa | gato | , 1 | 097 |
| 1045 | gtc | gcc | aa | ċàà | ġà | ċŧŧ | ttt | :tt | : é é | gå | ıģģ | aa | àċ | έŧ | 446 | ff | åξ | çŧ | ۵į, | lll | gate | 1 | 094 |
| 1098 | III | <u> </u> | CE | 779 | ga | ffi | P | ŦĠ | ta a | gą | ga | aa. | ą¢t | ţaa | 7 4 9 | 759 | :ąç | ġą | ţţ. | jąą | gact | 1 | 147 |
| 1148 | cca | tgc | cat | itc | ga | tct | àċ | tġ | ráá | ıġá | ġá | à à a | áct | Lga | ąέċ | įέċ | :dc | ģå | ξĘĠ | jaa | gat gaç gaç | : 1 | 144 |
| 1145 | | | | Ç | ag: | FF | rça | aa | aa | ֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֡֓֓֓֓֓֡֓֓֡֓֡֓֓֡֓ | Fa | F9. | aåq | jąc | 194 | aţ | gg | ŧ¢: | tţ | ţ | FFC | : 1 | 197 |
| 1198 | tati | ggc | att | ca | agı | cta | ca | aa • | ag | tt | ta | tġ. | àá¢ | jác | gģā | iát | φģ | f¢. | ata | ta | FFC | : 1 | 194 |
| 1195 | ĮŢŢ | ווון | ַלְלָב <u>ָ</u> | 11. | II. | וַן | Ï | <u>a</u> a | ŢŢ | ac T | 19 | age | rài | 99 | ga | aç | 77 | ågo | rat | FF | ggt ggt | 1 | 247 |
| 1248 | Ctac | :99 :20 | cas | 40 | 200 | | | aa .: | ac | ac | ag. | ago | ta a | làá | ıġá | áċ | ct | àġ | :at | : £ £ | ąątą | 1 | 244 |
| 1245 | CEG | ag. | | Ĭ | ֓֓֟֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓ | וַיִי | J. | ŢŢ | ij | 17 | ĮĮ. | aça | 199 | 161 | a t | 19 | 99 | ۴۴ | J F q | çq | ggtg atçá gtca | 1: | 297 |
| 1298 | ctat | ca | gaa | ac. | tcc | , y c | aa. | a. tå | -c | ay c+ | ta. | a ca | gg | ICT | at | tċ | ġġ: | tto | gta | ċġ | gtċa | 1: | 294 |
| 1295 | ctg | ca | III daa | II. | ΙIJ | I | ĬĬ | ĮŢ | IJ | ΪĮ | בַלַן. | [נֻיַן | 11 | | 75 | a a | ag: | ֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓ | 77 | ga | gtca agca agca | 1: | 347 |
| 1348 | agat | :aa | ctc | ta | cat | at | ati | αά | tt | 50 | ta | 720 | .aa | 20 | gc | aa | aga | 350 | ct | ġà | aġċà • | | 344 |
| 1345 | agat agat | :aa | Ff | fg | ge | at | <u>af</u> | | ίſ | ĬΪ | | ֓֞֞֞֞֞֞֞֓֞֞֞֓֓֓֞֞֩֞֓֓֞֞֩֞֓֓֓֓֞֓֡֓֞֡֓֡֓֞֡֓֡֡֡֡֡֡֡֡ | ֓֟֝֟֝֟֟֟֟֝֟֝֟֟֝֟֝֟֝֟֝֟֝֟֝ <u>֚֚֚֚֚֟</u> | וַלַ | ٢٩ | ĬĮ. | <u>Y</u> C | | Ţ | ac. | 7555 | | 397 |
| 1398 | caat | gça | ąąg | çaı | ‡gç | ąg | çad | gŧ | Çţ | tt | aac | ca | iåa | aa | aa | GG. | y Ca a a â | 390 | .aa | ac. | accc | | 394 |
| 1395 | caat | ģċ | lag | cat | Fåå | Ιdg | čá | Ĩξ | çΕ | ff | | ca | II | ď | I | | | ֓֓֞֟֓֓֓֟֝֟֓֓֟֝֟֓֓֓֟֝֟֓֓֟֝֟֓֓֟֝֟֓֓֓֟֝֟֝ <u>֚֚</u> | ĬĮ. | ֓֞֟֝֟֓֟֟֝֟֟֟֟֝֟֟֟֝֟֟֟֝֟֟֝֟֝֟֝֟֝֟֝֟֝֟֝֟֝֟ | | | 147 |
| 1448 | gcat | FF | 777 | ĊŦ | FFF | cg | gg1 | ţţ, | ąg | ągo | ;; | aţ | ŧα | ta | ca | aci | tad | at | at | ta | aga | | 97 |
| 1445 | ġċát | tċt | fff | ċŧŧ | FFF | tģ | gai | LŁa | ١ģ | ago | 744 | : d t | fţ | fģ | ĮΙ | | | I | IĮ | ĮĮ, | ֓֓֓֓֓֓֓֟֓֓֓֟֓֓֓֓֟֓֓֓֓֓֟֓֓֓֓֓֓֓֓֓֓֓֓֟֓֓֓֓ | | 94 |
| 1498 | aacc | aga | 146 | ġţţ | FFF | FF | 499 | aġ | 761 | FF | ? | çç | ġç | FF | gç(| - Çaa | act | ta | ga | tai | .aat | | 47 |
| 1495 | aacc | aġa | iáċ | att | ċċŧ | tt | àġá | aad | cct | ffe | ; £ 6 | ငှင် | ĝς | fξ | åς | caa | cf | ŧ₫ | ga | få | dat | | 44 |
| 1548 1545 | 9999 | ffi | 77 | 559 | 199 | F F' | 779 | à | rti | a t | ra t | FF | ŧŧ. | çg | ąţ | FFI | gţ | аţ | Fg. | FFI | ţçċ | | 97 |
| 1598 | 9999 | | .cc | tto | 199 | te | tto | ad | eta | ito | àt | ċŧ | ££. | ċġ | ato | ; to | :ģŧ | ct | ξģ | ff | fee | 15 | 94 |
| 1595 | atgt | | Ĭ | 111 | 75 | ŦŦ: | | JÇ(| Sec | 799 | 110 | 49 | Ça | Fg | age | jat | gŧ | Çą | F F | Fgt | gaġ | 16 | 47 |
| 1648 | acat | tta | ce | | .gc | | cgç | ici | 3CC | caa | ac | àġ | Ċā: | tġ. | aġ | ját | ġŧ | ċå | ¢ť. | tģŧ | gag | 16 | 44 |
| 1645 | acat | III | II. | וַוַן | 10 | | | | | 191 | 90 | ag | <u>የ</u> የ | a t | ggt | řtí | aa | gç | ትት | 769 | tçġ | 16 | 97 |
| 1698 | aaaa | ata | ati | | .ay → | ati | | | | 191 | gc | ag | t t | ato | ġġŧ | aa | ıga | ġċ | tt | àċċ | :¢ċġ | 16 | _ |
| 1695 | aaag aaag | ΪIJ | aE | | Į | ֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓ | ֚֡֓֓֞֟֝֟֝֓֓֓֓֟֝֟֓֓֓֓֓֟֝֓֓֓֟֝֓֓֓֓֟֡֓֓֓֓֓֟֡֓֓֓֓֓֡֡֡֡֡֓֓֓֓֓֡֡֡֡֓֓֓֓֡֡֡֡֓֓֡֡֡֡֓֡֡֡֡֡ | ֡֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓ | إو | ֓֓֓֓֓֓֟֟֓֓֓֓֟֟֓֓֓֓֓֟֓֓֓֓֓֓֟֓֓֓֓֓֓֓֓֓֓ | Ï | a E | 49 | [9] | C | F | tt | aa. | a Fi | 779 | tgā | 17 | |
| 1745 | ggat | cct | tc | ita | ta | tat | - 2 - | à t | · Ca | 100 | | C C | agı | cgi | | ca | cc | aa. | ata | ààà | tġà | 17 | |
| 1745 | ggat ggat | ççt | fc | raa | Į. | ΙΙ | | | וו | ŢŢ | | ַן דַ | ֓֞֞֞֓֓֓֞֟֓֓֓֟֝֟֓֓֓֟֝֟֓֓֟֝֟֓֓֟֝֓֓֟֝֟֓֓֓֟֝֓֓֓֟֝֓֓֓֟֝֓֡֝֡֓֡֝֡ | ֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֟֓֓֓֟֓֓֓֓֓֟֓֓֓֓֓֟֓֓֓֓֟֓֓֓֓ | ֓֞֟֓֓֓֓֓֓֓֓֓֓֓֓֓֓֟֟֓֓֓֓֓֓֓֓֓֓֓֓֡֓֡֓֓֓֓֡֓֡֓֡֓֡֓֡֡֓֡ | Ì | FF | <u>ff</u> | ata | ן בין | CFA | 17 | |
| 1795 | 474 | | 944 | 3 U L | aa | ccc | ıaa | aa | I T. T | ~~ | | ~ a | P 3/ | 73 | | | _ : | | | | • | 17 | |
| 1795 | atct | faf | ala | Ϊåf | aa. | ato | CC | ga | | tt | ga | 77 | | 700 | -0+ | ad. | إم | ֓֟֓֓֓֟֟֓֓֓֟֟֓֓֓֓֓֓֓֟֟֓֓֓֓֟֓֓֓֓֓֓֓֓֓֟֓֓֓֓ | aa. | 144 | aaa | 18 | |
| 1845 | · F | | | | • • | | | • • | • • | | | | | | | | | | | | - | 18 | |
| 1845 | tgåt | ata | aaç | jta | aa | ttt | at | tq | at | ac | aa | aad | gt: | o | | tr | | | • • • • • • | | | 18: | |
| | | | | | | | | - | | _ | _ | | , | , | | | | | | | 444 | 10 | 74 |

F 12

WO 99/63055

GAP program of Genetics Computer Group blosum62.cmp Gap Weight: Average Match: 2.912 12 Length Weight: Average Mismatch: -2.003 Quality: Length: 530 Ratio: 5.147 Gaps: 0 Percent Similarity: 99.623 Percent Identity: 99.057 Match display thresholds for the alignment(s): | = IDENTITY Adonis palaestina ε-cyclase #3 x Adonis palaestina ε-cyclase #5 1 MELLGVRNLISSCPVWTFGTRNLSSSKLAYNIHRYGSSCRVDFQVRADGG 50 1 MELLGVRNLISSCPVWTFGTRNLSSSKLAYNIHRYGSSCRVDFQVRADGG 50 51 SGSRTSVAYKEGFVDEEDFIKAGGSELLFVQMQQTKSMEKQAKLADKLPP 100 51 SGSRSSVAYKEGFVDEEDFIKAGGSELLFVQMQQTKSMEKQAKLADKLPP 100 101 IPFGESVMDLVVIGCGPAGLSLAAEAAKLGLKVGLIGPDLPFTNNYGVWE 150 101 IPFGESVMDLVVIGCGPAGLSLAAEAAKLGLKVGLIGPDLPFTNNYGVWE 150 151 DEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKR 200 151 DEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKR 200 201 CVESGVSYLNSKVERITEAGDGHSLVVCENDIFIPCRLATVASGAASGKL 250 201 CVESGVSYLDSKVERITEAGDGHSLVVCENEIFIPCRLATVASGAASGKL 250 251 LEYEVGGPRVCVQTAYGVEVEVENNPYDPNLMVFMDYRDYMQQKLQCSEE 300 251 LEYEVGGPRVCVQTAYGVEVEVENNPYDPNLMVFMDYRDYMQQKLQCSEE 300 301 EYPTFLYVMPMSPTRLFFEETCLASKDAMPFDLLKRKLMSRLKTLGIQVT 350 301 EYPTFLYVMPMSPTRLFFEETCLASKDAMPFDLLKRKLMSRLKTLGIQVT 350 351 KIYEEEWSYIPVGGSLPNTEQKNLAFGAAASMVHPATGYSVVRSLSEAPK 400 351 KVYEEEWSYIPVGGSLPNTEQKNLAFGAAASMVHPATGYSVVRSLSEAPK 400 401 YASVIAKILKQDNSAYVVSGQSSAVNISMQAWSSLWPKERKRQRAFFLFG 450 401 YASVIAKILKQDNSAYVVSGQSSAVNISMQAWSSLWPKERKRQRAFFLFG 450 451 LELIVQLDIEATRTFFRTFFRLPTWMWWGFLGSSLSSFDLVLFSMYMFVL 500 451 LELIVQLDIEATRTFFRTFFRLPTWMWWGFLGSSLSSFDLVLFSMYMFVL 500 501 APNSMRMSLVRHLLSDPSGAVMVKAYLER* 530 501 APNSMRMSLVRHLLSDPSGAVMVRAYLER* 530

FIGURE 24

÷2.

| Ċ | j | |
|---|---|--|
| G | a | |
| 2 | 5 | |
| ζ | | |
| | 7 | |
| | | |

| | - 222204240HBB07 | | |
|---------|--|--|--|
| | 1003 1003 1003 1004 1004 1004 1004 1004 | 205 205 204 204 204 204 204 181 181 182 182 183 184 185 | 161 312 311 311 311 316 298 298 289 289 289 |
| * | PPIS PPIS PPIP PROIS PFEL DFEL DFEL DFEL | 220 CVES CVES CVES CVES CVES CVES CVES CVES | |
| | | | |
| 100 | VVVRYSYRNIRFGL-CSVRASGGSSGSESCVAVREDFADEEDFVKAGGBELLFVGHGONRDHDEGSKIVDRISSKLAYNIHRYGSSCRVDPQVRADGGSGSRSSVAYRGFVDEEDFIKAGGBELLFVGHGONRDHDEGSKINDKISSKLAYNIHRYGSSCRVDPQVRADGGSGSRTSVAYRGFVDEEDFIKAGGBELLFVGHGOTKSHEKDAKIADKISSKLAYNIHRYGSSCRVDFQVRADGGSGSRTSVAYRGFVDEEDFIKAGGBELLFVGHGOTKSHEKDAKIADKIPBLLKORRFTNLSA-SSGLRQIKCSAKSDRCVVDKQQISVADEEDFYKAGGBELFVGHGOTKSHEKDAKISKLISKHINGGREGIFLAY-EQYEGKCNSSSGSSCVVDKEDFADEEDYVKAGGBELFVGHGOTKSHESGSKLSENSKOIKCNAAKSQLVVKQSIBEEDFYKAGGBELFVGHGOTKSHENDGSKENHSCHUKTPNKLDFFILHGFAEKQHLVSTSKLQNQVFRIASRNIHPCRNGTVKARGSALLLDLUVPETKKENH | * HEELLER HEELLER HEELLER HEELLER HEELLER HEELLER HEELLER KSKWOR | THE PARTY OF THE P |
| Ä | HERE HERE HERE HERE HELV HELV HELV HELV HELV HELV HELV HEL | HILL HILL WOLL WOOD WOOD WOOD WOOD WOOD | 320 LEAR LEAR SEEE SEEE LEAR LEAR LEAG RNSKI RNSKI RNSKI RNSKI RNSKI |
| | ONKI ONKE ONKE ONKE ONKE ONKE ONKE ONKE ONKE | CRVC CRVC CRVC CRVC CRVC CRVC CRVC CRVC | 320 2 VRHDAQSLBAK 2 TNEKVRSLBAE 2 HQOKLQCSBEE 2 MQOKLQCSBEE 2 SKHKPESLBAK 2 SKHKPESLBAK 2 TKHRQQSLBAK 2 TKHRQQSLBAK 3 TKHRQCSLBAK 3 |
| • | VOUC VOUC VOUC VOUC VOUC VOUC VOUC VOUC | | -VRH -TNE -MQQ -MQQ -SKH -TKH -TKH -TKH CNNV CNNV SNNW |
| | KAGGSELLE KAGGSE | DDDD FILES DDDD FILES DDDD FILES DDDD FILES DDD FILES DD FILES | SHIP SHIP SHIP SHIP SHIP SHIP SHIP SHIP |
| 80 | AGGBELL AGGBELL AGGBELL AGGBELL AGGBELL AGGBELL SRSVC RSVC RSVC RSVC ROVC KLGQKYC KKCRNGYI | TODOS ILE TODOS | |
| ∞ | Drvk Drvk Drvk Drvk Criv STL STL STL | | 300 |
| | -ADEEDFY -VDEEDFY SVADEEDTY SVADEEDTY -ADEEDTY -REFERSENTH -KKFCEGLG SKKFCETL SKKFCETL SKKFCETL TREGRKKSQI | WRETI WRETI WRDTI WRDTI WRDTI WRDTI WRDTI WRDTV SGAN SGAN SGAN | D CONK |
| • | TRRPPVVRRYSTRIFGL-CSVRASGGGSSGSBSCVAVREDFADEEDFVFAGGBELLFV TRNLSSSKLAYNIHRYGSSCRVDPQVRADGGSGSRSSVAYKGFVDEEDFIFAGGBELLFV TRNLSSSKLAYNIHRYGSSCRVDPQVRADGGSGSRSSVAYKGFVDEEDFIFAGGBELFFV TRNLSSSKLAYNIHRYGSSCRVDFQVRADGGSGSRTSVAYKGFVDEEDFIFAGGBELFFV TRHRPSLLKQRRFTHLSA-SSSLRQIKCSAKSDRCVVDKGGISVADBEDYVFAGGBLLFVMSGGELCQKSIFLAY-EQYESKCNSSSGSDSCVVDKEDFADEEDYVFAGGBLLFVMDTLLKTPNKLDFFIPQFHGFEEKQHLVSTSKLQVQVFRIASRNIHPCRNGTVKMDTLLKTPNKLELLPTLHGFAEKQHLVSTSKLQNQVFRIASRNIHPCRNGTVKMDTLLKTPNKLELLPTLHGFAAVKASTFRSEKHNFGSRKFCETLGRSVCVKMDTLLKTPNKLELLPTHGPAVKASTFRSEKHNFGSRKFCETLGRSVCVKMDTLLKTPNKLELLPTHRUGSVKASSFNSVKPHKFGSRKICENWGKUCVKMDTLLKTPNKLELLPTHRUGSVKASSFNSVKPHKFGSRKICENWGKUCVKMDTLLKTPNKLELLPTHRUGSKUNSSFNSVKPHKFGSRKICENWGKUCVKMDTLLKTPNKLELLYPLHELAKRHFLSPSPNPQNPNFKFFSRKPYQKKCRNGYIG | 180 GIEHW GIEHW GIEHW GLDTIN GLDTIN GLDTIN GLDATIN GLDATIN GLDATIN GLDATIN GLDATIN | * CONNOTE CONTOTE CONNOTE CONNOTE CONNOTE CONNOTE CONNOTE CONNOTE CONNOTE CONTOTE CONNOTE CONN |
| | AVRE AVRE AVRE ODE ODE OF OF OF OF OF OF OF OF OF OF OF OF OF | | |
| | RSCV/ RTSVI CV OSCVV PYHE FSSVF F | TROUGH BROUGH BROUGH FROUGH FRANDI BR | AYGV AYGV AYGIB AYGIB AYGIB AYGIB AYGIB AYGIB AYGIB AYGIB AYGIB |
| 9 | SSGS 3SGS 3SGS SSGS | PEERD I DEERD I | |
| | RKFPVVKRISTRNIRFGL-CSVRASGGGSSGSESCVAVREDF TRNLSSSKLAYNIHRYGSSCRVDFQVRADGGSGSRSSVAYKEGF TRNLSSSKLAYNIHRYGSSCRVDFQVRADGGSGSRSSVAYKEGF IRHKFSLLKQRRFTNLSA-SSSLRQIKCSAKSDRCVVDKQGI -WASGGELCQEKSIFLAY-EQYEGKCNSSGSSSCVVDKEDF | INYGVWP INYGVWP INYGVWP INYGVWP INYGVWP INYGVWP INYGVWP INYGVWP INYGVWP INYGVWP INYGVWP INYGVWP INYGVWP INYGVWP INYGVWP | THE TOTAL THE TO |
| • | - RAI DPQVI DPQVI IRQI IRPA IGPA- IGPS- IGPS- IGPS- IGPS- IGPS- IGPS- IGPS- IGPS- IGPS- IGPS- IGPS- IGPS- IGPS- IGPS- IGPS- | | |
| | SCRVI | TO THE PERSON OF | SHELLO SH |
| 04 | RKFPVVKRYSTRNIRFGL-CSVRAS TRNLSSSKLAINIHRYGSSCRVDFQVR TRNLSSSKLAINIHRYGSSCRVDFQVR TRHKFSLLKQRRTHLSA-SSSLRQIK KDTLLKTPNKLDFFIPQFHGFR HDTLLKTPNKLBFLPTLHGFA- | | |
| 4 | YRNI YNIH YNIH YNIH YRFTI YRFNI YTPNN (TPNN (TPNN | CS ID | |
| | KRYS SKLA SKLA LLKQI SLCQI OTLLI OTLLR TLLR | | CERTAIN CONTROL OF THE CONTROL OF TH |
| * | TEPAN TESSI | 140 MESTEL MESTEL MESTEL MOUSEA COVSEA COVSE | TOTAL |
| | GTTRE GTTRE STRE | TAGO MEMBERS STREET | 581111511111111 |
| | ATT. | TAO DAGLAHAABAKA DAGLAHAABAKA DAGLAHAABAKA DAGLAHAABAKA DAGLAHAABAKA DAGLAHAABAKA DAGLAHAABAKA DAGLAHAABAKA DAGLAHAABAKA DAGLAVAQOVSEA DAGLAVAQOVSEA DAGLAVAQOVSEA | VECEDD VACOBNO VACOBNO VACOBNO VECEDO |
| 20 | HECVGARNP-AAMAVSTPFSM9-CRRKFPVVRRYSYRNIRFGL-CSVRASGGGSSGSSCVAVREDFADEEFFV HELLGVRNLISSCFWT-FGTRNLSSSKLAYNIHRYGSSCRVDFQVRADGGGSRSSVAYKEGFVDEEFFT HELLGVRNLISSCFWT-FGTRNLSSSKLAYNIHRYGSSCRVDFQVRADGGSGSRTSVAYKEGFVDEEFFT HECKGARNMTATMAVFTCFRFTDCNIRHKFSLLKQRRFTNLSA-SSSLRQIKCSAKSDRCVVDKGGISVADEEFT HECVGVQNV-GAMAVLTRFRIMRMSGGELCQEKSIFLAY-EQYESKCNSSSGSDSCVVDKEDFADEEFYT HECVGVQNV | 120 130 140 16 | - 240 GVLTINSFUDRIVERTRIGHSTOPERDOWN GVSTISSRUDSTITATORISTIVA DDRAWN GVSTISSRUDSTITATORISTIVA DDRAWN GVSTISSRUSRITATORISTIVA GRAND FRI GVSTISSRUSRITATORISTIVA GRAND FRI GVSTISSRUSRITATORISTIVA GRAND TRI GVA FROAKUIKVIH - RESKSHLI GNDGITH VAR FROAKUIKVIH - RESKSHLI GNDGVTH VAR FROAKUIKVIH - RESKSHLI GNDGVTH OUT FROAKUIKVIH - RESKSTLIGGSDGVTH GVA FROAKUIK - RESKSTLIGGSDGVTH GVA FROAKUIK - RESKSTLIGGSDGVTH GVA FROAKUIK - RESKSTLIGGSDGVTH GVA FROAKUIK |
| | HAAVE | BEAVY BILLYVI | H - H - H - H - H - H - H - H - H - H - |
| • | HECYGARNP - AAMAN HELLGYRNL - I HELLGYRNL - I HECYGARNHTATHAN HECYGYGNY - GAHAN HERAG - HHTATHAN | 120 DOALD BESVED | LTINSTORTORTER STIDSTORTER STI |
| | SCUGARA SLLGVRA SLLGVRA SCUGAR | 12 | LTINSKY STIDSKY STIDSKY STIDSKY STIDSKY KHOAKY KPHOAKY KPHOAKY KPHOAKY KPHOAKY |
| į | | 120 IG | GVLTINSKY GVSTIDSKY GVSTIDSKY GVSTIDSKY GVLTINSKY GVKTHORKY GVKTHORKY GVKTHORKY GVKTHORKY GVKTHORKY GVKTHORKY GVKTHORKY |
| - | = = = = = = = = = = = = = = = = = = = | 00 00 00 00 00 00 00 00 00 00 00 00 00 | |
| | Arabidopsiss Adoniss1 Adoniss2 Lettucess Tomatos Arabidopsiss Adoniss Peppers Tomatos Tobaccos Marigolds | PotatoB Arabidopsiss Adonissi Adonissi Adonissi Tomatos Amarigolds Arabidopsiss Peppers Peppers Comatos Arigolds | |
| PotatoB | Arabidops AdonisEl AdonisE2 LettuceEE TomatoE AdonisB PepperB PopacB TomatoB TobaccoB MarigoldB | PotatoB Arabidops AdonisE1 AdonisE1 LettuceER MarigoldB Arabidops: PepperB TomatoB TomatoB TobaccoB MarigoldB | Potatos Arabidopsiss Adonissi Adonissi Lettucess Lettucess Arabidopsiss Arabidopsiss Arabidopsiss Peppers Peppers Comatos Peppers |
| Pot | Are Add Add Tour Are Are Are Are Tour Tour Tour | PotatoB Arabido Adoniss Adoniss Pertuce Adoniss PepperB Tobaccoi Marigolo | Potatos Arabidopsiss Adonissi Adonissi Lettucess Tomatos Arabidopsiss Adoniss Peppers Tomatos Tobaccos Marigolds |
| | | | |

| _ | |
|-------|--------|
| | • |
| + | J |
| 4000 | : |
| ć | ָ כ |
| ` | |
| ሪ |) |
| G | 1 |
| 2 | ξ |
| TCITO | 5 |
| ۲ | 4 |
| G | 4 |

| 27 7 30 | # ¹ 2. |
|---|--|
| 2 4 4 4 4 4 4 4 2 2 2 2 2 4 4 4 4 4 4 4 | 8 9 10 0 0 0 0 11 0 0 0 0 0 0 0 0 0 0 0 0 |
| 11753 : NS: : NS: : VSGQ : : VSGQ : : LLTS- : | |
| | 540 THATLIRIYLIF- THATINIKY SHAVINA KILER- SHAVINA KILER- SHAVINIKA KILI THATILIRIYLIF THATILIRIYLIF LUVINI INILLODE |
| FYLLANT RONHSK SYLANT RODNSA SYLANT RODNSA SYLANT RODOSK SYLANT ROHYSK SYLANT ROHYSK SYLANT ROHYSK SYLANT ROHYSK SYLANT ROHYSK NAIL RYTH ROSERD NAIL RYTH ROSERD NSIVQY BUSERSH NAIL RYTH ROSERD NSIVQY BUSERSH NSIVQY BUSERSH NSIVQY BUSERSH NSIVQY BUSERSH NSIVQY BUSERSH NSIVQY BUSERSH NSIVQY BUSERSH | 540 SGATH SGATH SGATH SGATH |
| FYLLAN SSCIPT SSCIPT SSCIPT WAILE WAILE WAILE WAILE | |
| APPKY | * 540 ** 520 ** 540 ** 520 ** 540 ** 520 ** 540 ** |
| | 520 TIANDHRRGIIR TVISHNICHRRGIIR TVISHNICHRRGIIN TVIANSHRHSHVR TVIANDHRRGIIV TIANDHRRGIIV TIANDHRRITH TIANDHRRGIIV TIANDHRRGIIV TIANDHRRGIIV |
| | 520 STATE OF THE STATE OF THE S |
| MAND STORY ON WHE STORY ON W | MIGAPES MIG |
| | A CACACACACACACACACACACACACACACACACACAC |
| A CANADA | |
| | |
| | |
| | |
| HF>HFHHHHHHHH | 480 TEGERS TEGERS TEGERS TEGERS TEGERS TEGERS TEGERS |
| RINTIGURIKE RETTIGIOUTE RETTIG | THE PROPERTY OF THE PROPERTY O |
| STANDARD STANDARD | |
| | |
| ASKDAMETDILITELIA ASKDAMETDILITELIA ASKDAMETDILITELIA ASKDAMETDILITELIA ASKDAMETDILITELIA VARPGIRNEDIGEN VARPGIRNEDIGEN VARPGIRNEDIGEN VARPGIRNEDIGEN VARPGIRNEDIGEN VARPGIRNEDIGEN | |
| | STES-ESTEANTHWE SSAVNISHOLMSSEND SSAVNISHOLMSSEND SSINSISH SENDEN KYT-NISKOMMTHWE RYTTNISKOMMTHWE LEGDELSARVWED GWE -SGDELSARVWED GWE -LONELSARVWED HWE -LONELSARVWED HWE -LONELSARVWED HWE -LONELSARVWED HWE -LONELSARVWED HWE -LONELSARVWED HWE -LONELSARVWED HWE -SGNDLAADVWED HWE |
| SPTRVSESSION TKSRLFGESTOC BPTRLFGESTOC SPTRVFGESTOC SPTRVFGESTOC SPTRVFGESTOC SPTRVFGESTOC SPTRVFGESTOC SPTRVFGESTOC SPTRVFGESTOC SPTRVFGESTOC SPTRVFGESTOC SPTRFGESTOC SPTRFG | STES-ESTONE SSAVNISHONE SSAVNISHONE SSAVNISHONE SSAVNISHONE SSATTANE LEGDELSARV -SGDELSARV -LONGLESARV -LONGLESARV -LONGLESARV -SGDELSARV -SGDELSARV |
| SSNALL OF PROPERTY | STPS-FIST SSAVNESS SSAVNESS STREETS ERGDGLSA SGDBLSA - SGDBLSA - LUNBLSA - SGNDLAA |
| | |
| Potatos Arabidopsiss Adonissi Adonissi Tomatos Arabidopsiss Adoniss Peppers Tomatos Tobaccos | Potatos Arabidopsies Adonissi Adonissz Lettucess Tomatos Arabidopsies Adoniss Peppers Tomatos Tobaccos |
| Pol Add Add Mar Add Tom Mar Daf | Ars Pot Tour Tour Part Part Part Part Part Part Part Par |
| | |

 $\sigma_{\omega}^{i_{\omega}}$

FIGURE 26

| | 00 00 00 00 00 00 00 00 00 00 00 00 00 | 57 008 07 112 05 | LBLLU84 | 9 11 11 14 4 | · |
|-----|---|--|--|--|--|
| | 1 103 1 102 1 102 1 100 1 100 | 208 209 207 207 207 202 1205 | 1167 1318 1317 1317 1315 1304 | . 276 . 427 . 427 . 427 . 424 | |
| ٠ | TOTAL PPIS TADKIPPIP TADKIPPIP TSEKLAGIP TSEKRAGIS TSDEKROIS | V V V V V V V V V V V V V V V V V V V | ************************************** | 440 TPS-II NI SAVNI SAVNI YT-NI | 33 5524 5529 533 516 |
| | SKLYDKLPPIS AKLADKLPPIP AKLADKLPPIP SKLSDEKLAQIP SKLSDEKRAQIP SKLSDEKRAQIP | E E E E E E | | filtsss: NS VSGQS: VSGQS: NSGGS: IISLGK) ILTS-S: | 00 00 00 00 00 00 |
| 100 | | ARLINISSINA RRICHISSINA RRICHISSINA RRICHISSINA RRICHISSINA ROLINISSINA | 320 SARYPISM SEEVERS SEEVERS SARYPISM SARYPISM SARYPISM | * 440 TRONHSKNMLTSSSTPS-I TREETTKQINSN TRODNSAYVVSGOSSAVN TRODGSKEMISLGKYT-N TRODGSKEMISLGKYT-N TROHYSKNMLTS-SSIPSI | 540 GATLIRTITE GATHIRTITE GAVURALIER GATHURALIER GATHURALIER GATHURALIER |
| | FVQMQON FVQMQOT FVQMQOT FVQMQOK FVQMQOK | SOSTOS | | | DPTGAT DPTGAT DPSGAV DPSGAV DPTGAT DPTGAT |
| | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 | 200 PILJGRAYGRV PILJGRAYGRV PVLJGRAYGRV PILJGRAYGRV PILJGRAYGRV | TYRHDAGSL STANGKURSL STMQQKLQCS STMQQKLQCS STKHKPESCL STKHKSQSL | KCO FVI KYN SVI KYN SVI KYN SVI KYN SVI KYN SVI KYN SVI | IRHILIS INHILIS VRHILIS VRHILIS IRHILIS VRHILIS |
| 80 | 155555 | VXLDODD P VYLDDD RP VYLD WA P VYLDD ND P VYLDDD S P VYLDDD S P | 300 SLAVISADIAD VYREDAGSL DOMINISTORED Y TNEKVRSL NLAVISADIAD YRDGYCCS NLAVISADIAD SKRKPESCS SLAVISADIAD SKRKPESCS SLAVISADIAD Y TKHKSGSL | CSLSEAL CSLSEAL CSLSEAL CSLSEAL CSLSEAL CSLSEAL CSLSEAL | 520 PNDARRGI PNNIRKGI PNSARASI PNSIRASI PHSIRASI PNDARGI |
| | | REGIVE REGIVE REGIVE REGIVE REGIVE | WIS 400 WIND OF STANDOOR OF ST | GYSVVF GYSVVF GYSVVF GYSVVF GYSVVF | I TAPA |
| • | KFPVVKRYSYRNIRFGL-CSVRASGGGSSGSBSCVAVREDFAD RNLSSSKLAYNIHRYGSSCRVDFQVRADGGSGSRSGVAYKEGFAD RNLSSSKLAYNIHRYGSSCRVDFQVRADGGSGSRTSVAYKEGFVD RHKFSLLKQRRFTNLSA-SSSLRQIKCSAKSDRCVVDKQGISVAD RNSGGRLCQEKSIFLAY-EQTESKCNSSSGSDSCVVDKEDFAD | 180 CONTRACTOR | EVEVENNE EVENNE EVEVENNE EVEVENNE EVEVENNE EVEVENNE EVEVENNE EVEVENNE EVEVENNE EVEVENNE EVEVENNE EVEVENNE EVEVENNE EVEVENNE EV | LAFGAAASWVHPATGYSVVRSLSEAP LAFGAAASWVHPATGYSVVRSLSEAP LAFGAAASWVHPATGYSVVRSLSEAP LAFGAAASWVHPATGYSVVRSLSEAP LAFGAAASWVHPATGYSVVRSLSEAP LAFGAAASWVHPATGYSVVRSLSEAP | A A B B B A A A |
| 6 | RASGGGSSGSRSCVAVREDF DPQVRADGGSGSRSSVAYKEGF LRQIKCSAKSDRCVVDKQGI RSKCNSSGSDSCVVDKEDF KQIKCNAAKSQLVVKQEI | S S H H S S B | 5 5 5 5 H 5 H | AFGRANS AFGRANS AFGRANS AFGRANS AFGRANS AFGRANS | SXADI TSGDI TSGDI SSFDI SSTDI SSADI SSTDI SSTDI |
| 9 | RADGGS RADGGS RADGGS RCSAKSI KCNAAK | 100 PER | 280 RVSVQTAXGO RVCVQTAXGO RVCVQTAXGO RVCVQTAXGO RVCVQTAXGO RVCVQTAXGO | | MWGFLGSSL WWGFLGSSL WWGFLGSSL WWGFLGSSL WWGFLGSSL WWGFLGSSL |
| • | SVRA RVDFQV RVDFQV SSLRQI QYRSI | 160 PERMOTOLINATORNEDERNOGIO GPOLPFINNYGWNEDERNOGIO GPOLPFINNYGWNEDERNOGIO GPOLPFINNYGWNEDERNOGIO GPOLPFINNYGWNEDERNOGIO GPOLPFINNYGWNEDERNOGIO GPOLPFINNYGWNEDERNOGIO | | 380 GGSLPNT GGSLPNT GGSLPNT GGSLPNT GGSLPNT GGSLPNT | THE PROPERTY OF STREET |
| | YRNIRFGL-C YNIHRYGSGC YNIHRYGSGC RRYTHLSA-S REKSIFLAY-B | GPDLPF GPDLPF GPDLPF GPDLPF GPDLPF GPDLPF | | * *********************************** | A PERVI |
| 0 | LAYNIH LAYNIH LAYNIH KORRFI CORKSI | | 260 TVASGAASGKI TVASGAASGKI TVASGAASGKI TVASGAASGKI TVASGAASGKI TVASGAASGKI TVASGAASGKI | GVRIKEI KEEBBMSYIPVGGSLPNTEQKA GIRLKAYEEBBMSYIPVGGSLPNTEQKA GIQVIKYEEBBMSYIPVGGSLPNTEQKA GIQVIKYEEBBMSYIPVGGSLPNTEQKA GIRLTRAYEEBMSYIPVGGSLPNTEQKA GVRIKKIYEEBMSYIPVGGSLPNTEQKA | 480 HRESES HRI |
| • | indsser indsser indsser ingggre | 40 AESAKLGL AEDAKLGL AEBAKLGL AESAKLGL AESAKLGL AESAKLGL | PCREAT PCREAT PCREAT PCREAT PCREAT | 11GYRIKEIYEEEMSYIPYGGSLPNTEQKA TIGIRILKAYEEEMSYIPYGGSLPNTEQKA TIGIQYTKIYEEEMSYIPYGGSLPNTEQKA TIGIQYTKIYEEEMSYIPYGGSLPNTEQKA VY GIRITRAYEEEMSYIPYGGSLPNTEQKA TIGYRIKEIYEEEMSYIPYGGSLPNTEQKA | |
| | | AGENTANA OF THE PARTY OF THE PA | | FEREREE | SEFFECTION OF SE |
| 20 | VSTPST ISSOPVI IPTOPRI //ITROPRI | 120 -DGALDHVVIGGGPAGIALA -ESVVDIAVVIGGGPAGISLA -ESVVOLAVVIGGGPAGISLA -NCILDIAVVIGGGPAGIALA -QTVLDIAVVIGGGPAGIALA DSNCILDIAVVIGGGPAGIALA | 240 WENTY CHANGE CODYN THE CONTROL OF CONTROL THE CONTROL CHANGE IN THE CONTROL THE CONTROL CHANGE IN THE CONTROL OF CONTROL CHANGE CONTROL CHANGE CODYN THE CONTROL CONTROL CHANGE CONTROL CHANGE CONTROL CHANGE CONTROL CHANGE CONTROL CHANGE CONTROL CHANGE C | MPFD LLKWKL MPFD LLKWK MPFD LLKWK | 460 ORAFEL ORAFEL ORAFEL ORAFEL ORAFEL ORAFEL |
| | F-AAKAN G1 GTATKAN F-GAKAN | | VEATIVO TEASDO TEASDO TEASDO VEATIVO VEATIVO | DAMPED DAMPED DAMPED DAMPED BAMPED BAMPED BAMPED BAMPED BAMPED | POERKR PPERKR PKERKR PLERKR PLERKR PLERKR |
| • | Wecvgarnf-aamavstfess-crrffpvyrrystriffgl-csv Wellgyrnlisscpwt-fotrnlsssklatnihrygscrv Wellgyrnlisscpwt-fotrnlsssklatnihrygscrv Wecpgarnhtatmavftcerftdcnirhrsllkorrytnlsa-sss Wecvgvgnv-gamavlfrennrwsggelcorsifla | 120 • | * YDSKVDRI YDSKVBRI YDSKVBRI YDSKVBRI YDSKVBRI YDSKVBRI YDSKVBRI YDSKVBRI | ************************************** | 460 STOANWILMPOERKRORSFELFGLALIN SROANGILMPERKRORAFELFGLELIN SROANGSLAPFERKRORAFELFGLELIN SROANGSLAPFERKRORAFELFGLELIN SROANFLAMPLERKRORAFELFGLELIN STOANNILMPOERKRORSFFLFGLSHIU STOANNILMPOERKRORSFFLFGLALIU |
| | | | | | |
| | | 84 85 75 80 80 80 80 80 80 80 80 80 80 80 80 80 | M W W | 8 | ## ** |
| | Potatos Arabidopsiss Adonissi Lettucess Tomatos | tos idopi issi issi icesi tos | tos Idopi Issi Issi Icessi Iolds | COR Idopa SEST SEST SEST COR COR | 01 8 8 1 8 8 0 0 0 0 0 0 0 0 0 0 0 0 0 0 |
| | Potatos Arabidops Adonissi Adonissi Lettucess Tomatos Marigolds | Potatos Arabidopsiss Adonissi Adonissi Lettucess Tomatos | Potatos Arabidopsiss Adonissi Lettucess Tomatos | Potatos Arabidopsiss Adonissi Adonissi Lettucess Tomatos | Potatos Arabidopsiss Adonissi Lettucess Tomatos |
| | | | | | |

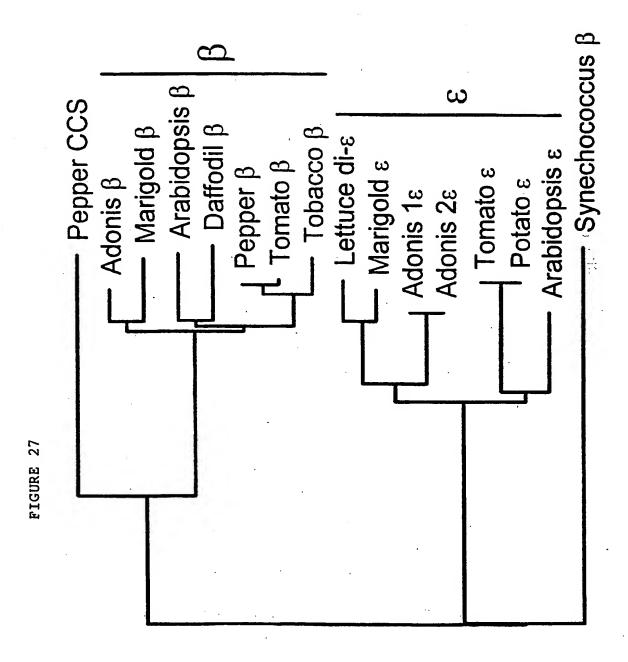




FIGURE 28

GAP of: Arabidopsis epsilon cyclase to Lettuce epsilon cyclase

Gap Weight: 12 Average Match: 2.912
Length Weight: 4 Average Mismatch: -2.003

Quality: 1837 Ratio: 3.499

Length: 534 Gaps:

Percent Similarity: 76.381 Percent Identity: 69.905

Match display thresholds for the alignment(s):

= IDENTITY

: = 2

1

Arabidopsis x Lettuce

| 1 MECVGARNF.AAMAVSTFPSWSCRRKFPVVKRYSYRNIRFGLCSVRA 46 : 1 MECFGARNMTATMAVFTCPRFTDCNIRHKFSLLKQRRFTNLSASSSLRQI 50 |
|--|
| |
| 47 SGGGSSGSESCVAVREDFADEEDFVKAGGSEILFVQMQQNKDMDEQSKLV 96 |
| 51 KCSAKSDRCVVDKQGISVADEEDYVKAGGSELFFVQMQRTKSMESQSKLS 100 |
| 97 DKLPPISIGDGALDHVVIGCGPAGLALAAESAKLGLKVGLIGPDLPFTNN 146 |
| : . |
| 147 YGVWEDEFNDLGLOKCIEHVWRETIVYLDDDKPITIGRAYGRVSRRLLHE 196 |
| : : :: : |
| 197 ELLRRCVESGVSYLSSKVDSITEASDGLRLVACDDNNVIPCRLATVASGA 246 |
| |
| 247 ASGKLLQYEVGGPRVCVQTAYGVEVEVENSPYDPDQMVFMDYRDYTNEKV 296 |
| |
| 297 RSLEAEYPTFLYAMPMTKSDLEFFFFTGLASKFURGER |
| . :: :: . : 346 301 ESLEAKYPTFLYVMAMSPTKIFFEETCLASREAMPFNLLKSKLMSRLKAM 350 |
| 347 GIRILKTYEEEWSYT DUGGGI DNTTROVAN A FRANCE . |
| : |
| 111 VGGSDFN1EQKNLAFGAAASMVHPATGYSVVRSL 400 |
| 397 SEAPKYASVIAEILREETTKQINSNISRQAWDTLWPPERKRQRAF 441 |
| |
| 442 FLFGLALIVQFDTEGIRSFFRTFFRLPKWMWOGFLGGTUTGGTUTGT |
| . . |
| 492 MFVISPNNLRKGLINHLISDPTGATMIKTVLKUL 525 |
| . : : : : 501 MFVIAPHSLRMELVRHLLSDPTGATMVKAYLTI* 534 |

SEQUENCE LISTING

| <110> | SUN, 2 | | | , FF | (ANC) | IS X. | • | | | | | * | | | |
|----------------------------------|----------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|-----|
| <120> | GENES METHOI | | | | | | THES | SIS F | AND 1 | IETAE | BOLIS | IA ME | ND | | |
| <130> | 8172-9 | 9023 | | | | | | | | | | | | ٠ | |
| <140> <141> | | | | IED | | | | ٠ | | | | | | | |
| <150> <151> | | | | | | | | | , | | | | | | |
| <150> <151> | | - | | | | | | | | | | | | | |
| <160> | 61 | | | | | | | | | | | | | | |
| <170> | Patent | In V | er. | 2.0 | | | | | | | | | | | |
| <210> <211> <212> <213> | 1860 DNA | lopsi | s th | nalia | ına | | | | | | | | | | |
| <220> <221> <222> | | (16 | 580) | | | | | | | | | | ` | | |
| <400> ACAAAA | _ | raat <i>e</i> | TAG | AT TO | CTCI | rttci | r GCT | rtgci | ATA | CCTI | · 『GATA | AGA A | ACAAT | TATAAC | 60 |
| AATGGT | GTAA (| STCTT | CTC | SC TO | STATT | rcga <i>i</i> | A ATT | ratt1 | rgga | GGAG | GAA <i>i</i> | Met | | TGT Cys | 117 |
| GTT GG Val Gl | G GCT y Ala 5 | AGG Arg | AAT Asn | TTC Phe | GCA Ala 10 | GCA Ala | ATG Met | GCG Ala | GTT Val | TCA Ser 15 | ACA Thr | TTT Phe | CCG Pro | TCA Ser | 165 |
| rgg Ag Irp Se 20 | T TGT r Cys | CGA Arg | AGG Arg | AAA Lys 25 | TTT Phe | CCA Pro | GTG Val | GTT Val | AAG Lys 30 | AGA Arg | TAC Tyr | AGC Ser | TAT Tyr | AGG Arg 35 | 213 |
| AAT AT Asn Il | T CGT e Arg | TTC | GGT Gly 40 | TTG Leu | TGT Cys | AGT Ser | GTC Val | AGA Arg 45 | GCT Ala | AGC Ser | GGC Gly | GGC Gly | GGA Gly 50 | AGT Ser | 261 |
| TCC GG Ser Gl | T AGT y Ser | GAG Glu 55 | AGT Ser | TGT Cys | GTA Val | GCG Ala | GTG Val 60 | AGA 'Arg | GAA Glu | GAT Asp | TTC Phe | GCT Ala 65 | GAC Asp | GAA Glu | 309 |
| GAA GA Glu As | T TTT p Phe 70 | GTG Val | AAA Lys | GCT Ala | GGT Gly | GGT Gly 75 | TCT Ser | GAG Glu | ATT Ile | CTA Leu | TTT Phe 80 | GTT Val | CAA Gln | ATG Met | 357 |
| CAG CA | G AAC | AAA | GAT | ATG | GAT | GAA | CAG | TCT | AAG | CTT | GTT | GAT | AAG | TTG | 405 |

| | W(| 99/6 | 3055 | | | | | | | | | | | | | PCT/US99/12121 |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|----------------|
| Gln | Gln 85 | Asn | Lys | Asp | Met | Asp 90 | Glu | Gln | Ser | Lys | Leu 95 | Val | Asp | Lys | Leu | .· <u>.</u> |
| CCT Pro 100 | Pro | ATA Ile | TCA Ser | ATT Ile | GGT Gly 105 | GAT Asp | GGT Gly | GCT Ala | TTG Leu | GAT Asp 110 | CAT His | GTG Val | GTT Val | ATT Ile | GGT Gly 115 | 453 . |
| TGT Cys | GGT Gly | CCT Pro | GCT Ala | GGT Gly 120 | TTA Leu | GCC Ala | TTG Leu | GCT Ala | GCA Ala 125 | GAA Glu | TCA Ser | GCT Ala | AAG Lys | CTT Leu 130 | GGA Gly | 501 |
| TTA Leu | AAA Lys | GTT Val | GGA Gly 135 | CTC Leu | ATT Ile | GGT Gly | CCA Pro | GAT Asp 140 | CTT Leu | CCT Pro | TTT Phe | ACT Thr | AAC Asn 145 | AAT Asn | TAC Tyr | 549 |
| GGT Gly | GTT Val | TGG Trp 150 | GAA Glu | GAT Asp | GAA Glu | TTC Phe | AAT Asn 155 | GAT Asp | CTT Leu | GGG Gly | CTG Leu | CAA Gln 160 | AAA Lys | TGT Cys | ATT Ile | 597 |
| GAG Glu | CAT His 165 | GTT Val | TGG Trp | AGA Arg | GAG Glu | ACT Thr 170 | ATT Ile | GTG Val | TAT Tyr | CTG Leu | GAT Asp 175 | GAT Asp | GAC Asp | AAG Lys | CCT Pro | 645 |
| ATT Ile 180 | ACC Thr | ATT Ile | GGC Gly | CGT Arg | GCT Ala 185 | TAT Tyr | GGA Gly | AGA Arg | GTT Val | AGT Ser 190 | CGA Arg | CGT Arg | TTG Leu | CTC Leu | CAT His 195 | 693 |
| GAG Glu | GAG Glu | CTT Leu | TTG Leu | AGG Arg 200 | AGG Arg | TGT Cys | GTC Val | GAG Glu | TCA Ser 205 | GGT Gly | GTC Val | TCG Ser | TAC Tyr | CTT Leu 210 | AGC Ser | 741 |
| TCG Ser | AAA Lys | GTT Val | GAC Asp 215 | AGC Ser | ATA Ile | ACA Thr | GAA Glu | GCT Ala 220 | TCT Ser | GAT Asp | GGC Gly | CTT Leu | AGA Arg 225 | CTT Leu | GTT Val | 789 |
| GCT Ala | TGT Cys | GAC Asp 230 | GAC Asp | AAT Asn | AAC Asn | GTC Val | ATT Ile 235 | CCC Pro | TGC Cys | AGG Arg | CTT Leu | GCC Ala 240 | ACT Thr | GTT Val | GCT Ala | 837 |
| TCT Ser | GGA Gly 245 | GCA Ala | GCT Ala | TCG Ser | GGA Gly | AAG Lys 250 | CTC Leu | TTG Leu | CAA Gln | TAC Tyr | GAA Glu 255 | GTT Val | GGT Gly | GGA Gly | CCT Pro | 885 |
| AGA Arg 260 | GTC Val | TGT Cys | GTG Val | CAA Gln | ACT Thr 265 | GCA Ala | TAC Tyr | GGC Gly | GTG Val | GAG Glu 270 | GTT Val | GAG Glu | GTG Val | GAA Glu | AAT Asn 275 | 933 |
| AGT Ser | CCA Pro | TAT Tyr | GAT Asp | CCA Pro 280 | GAT Asp | CAA Gln | ATG Met | GTT Val | TTC Phe 285 | ATG Met | GAT Asp | TAC Tyr | AGA Arg | GAT Asp 290 | TAT Tyr | 981 |
| ACT Thr | AAC Asn | GAG Glu | AAA Lys 295 | GTT Val | CGG Arg | AGC Ser | TTA Leu | GAA Glu 300 | GCT Ala | GAG Glu | TAT Tyr | CCA Pro | ACG Thr 305 | TTT Phe | CTG Leu | 1029 |
| TAC Tyr | GCC Ala | ATG Met 310 | CCT Pro | ATG Met | ACA Thr | Lys | TCA Ser 315 | AGA Arg | CTC Leu | TTC Phe | TTC Phe | GAG Glu 320 | GAG Glu | ACA Thr | TGT Cys | 1077 |

1125

TTG GCC TCA AAA GAT GTC ATG CCC TTT GAT TTG CTA AAA ACG AAG CTC Leu Ala Ser Lys Asp Val Met Pro Phe Asp Leu Leu Lys Thr Lys Leu

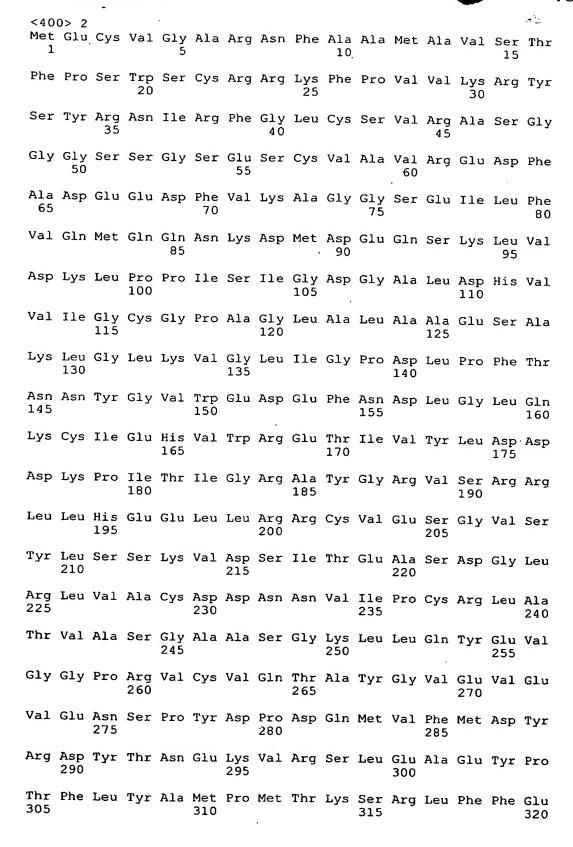
PCT/US99/12121

| WO 99/63055 | | |
|-------------|--|--|
| | | |

| | | - | | | | | | | | | | | | | | |
|-------------------|-------------------|------------|-------------------|------------|-------------------|------------|------------|-------------------|------------|-------------------|------------|------------|-------------------|------------|-------------------|--------|
| | 325 | | | | | 330 | | | | | 335 | | | | er iz | |
| ATG Met 340 | TTA Leu | AGA Arg | TTA Leu | GAT Asp | ACA Thr 345 | CTC Leu | GGA Gly | ATT Ile | CGA Arg | ATT Ile 350 | CTA Leu | AAG Lys | ACT Thr | TAC Tyr | GAA Glu 355 | 1173 |
| | GAG Glu | | | | | | | | | | | | | | | 1221 |
| | AAG Lys | | | | | | | | | | | | | | | 1269 . |
| | GGC Gly | | | | | | | | | | | | | | | 1317 |
| | GTC Val 405 | | | | | | | | | | | | | | | 1365 |
| | AAT Asn | | | | | | | | | | | | | | | 1413 |
| | AGA Arg | | | | | | | | | | | | | | | 1461 |
| TTC Phe | GAT Asp | ACC Thr | GAA Glu 455 | GGC Gly | ATT Ile | AGA Arg | AGC Ser | TTC Phe 460 | TTC Phe | CGT Arg | ACT Thr | TTC Phe | TTC Phe 465 | CGC Arg | CTT Leu | 1509 |
| | | | | | | | | | | | | | | | GGA Gly | 1557 |
| | CTC Leu 485 | | | | | | | | | | | Ser | | | AAT Asn | 1605 |
| | Arg | | | | | Asn | | | | | Asp | | | | GCA Ala 515 | 1653 |
| | ATG Met | | | | Tyr | | | | | ATTT | CTT. | ATCA | ACTC | TT | | 1700 |
| AGG | TTTG | TGT | TATA | TATA' | GT T | GATT | TATC | T GA | ATAA | TCGA | TCA | AAGA | ATG | GTAT | GTGGGT | 1760 |
| TAC | TAGG | AAG | TTGG | AAAC | AA A | CATG | TATA | G AA | TCTA | AGGA | GTG | SATCG | AAA | TGGA | GATGGA | 1820 |
| AAC | GAAA | AGA | AAAA | AATC | AG I | CTTT | GTTI | T GI | GGTT | 'AGTG | ; | | | | | 1860 |

<210> 2 <211> 524 <212> PRT <213> Arabidopsis thaliana

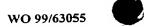




PCT/US99/12121

420

480



| | | - | | | | | | | | | | | | • | | |
|------------|---------------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-----|
| Glu | Thr | Cys | Leu | Ala 325 | Ser | Lys | Asp | Val | Met 330 | Pro | Phe | Asp | Leu | Leu 335 | Lys | |
| Thr | Lys | Leu | Met 340 | Leu : | Arg | Leu | Asp | Thr 345 | Leu | Gly | Ile | Arg | Ile 350 | Leu | Lys | |
| Thr | Tyr | Glu 355 | Glu | Glu | Trp | Ser | Tyr 360 | Ile | Pro | Val | Gly | Gly 365 | Ser | Leu | Pro | |
| Asn | Thr 370 | Glu | Gln | Lys | Asn | Leu 375 | Ala | Phe | Gly | Ala | Ala 380 | Ala | Ser | Met | Val | |
| His 385 | Pro | Ala | Thr | Gly | Tyr 390 | Ser | Val | Val | Arg | Ser 395 | Leu | Ser | Glu | Ala | Pro 400 | |
| Lys | Tyr | Ala | Ser | Val 405 | Ile | Ala | Glu | Ile | Leu 410 | Arg | Glu | Glu | Thr | Thr 415 | Lys | |
| Gln | Ile | Asn | Ser 420 | Asn | Ile | Ser | Arg | Gln 425 | Ala | Trp | Asp | Thr | Leu 430 | Trp | Pro | |
| Pro | Glu | Arg 435 | Lys | Arg | Gln | Arg | Ala 440 | Phe | Phe | Leu | Phe | Gly 445 | Leu | Ala | Leu | |
| Ile | Val 450 | Gln | Phe | Asp | Thr | Glu 455 | Gly | Ile | Arg | Ser | Phe 460 | Phe | Arg | Thr | Phe | |
| Phe 465 | Arg | Leu | Pro | Lys | Trp 470 | Met | Trp | Gln | Gly | Phe 475 | Leu | Gly | Ser | Thr | Leu 480 | |
| Thr | Ser | Gly | Asp | Leu 485 | Val | Leu | Phe | Ala | Leu 490 | Tyr | Met | Phe | Val | Ile 495 | | |
| Pro | Asn | Asn | Leu 500 | Arg | Lys | Gly | Leu | Ile 505 | Asn | His | Leu | Ile | Ser 510 | Asp | Pro | |
| Thr | Gly | Ala 515 | Thr | Met | Ile | Lys | Thr 520 | Tyr | Leu | Lys | Val | | | | | |
| <21: | 0> 3 l> 9! 2> Di 3> A: | AK | dops | is t | hali | ana | | | | | | | | | | |
| | 0> 3 CTTT(| CTC (| CTCC | тсст | CT A | CCGA | TTTC | C GA | CTCC | GCCT | ccc | GAAA | TCC | TTAT | CCGGAT | 60 |
| TCT | CTCC | GTC | TCTT | CGAT | TT A | AACG | CTTT | т ст | GTCT | GTTA | CGT | CGTC | GAA | GAAC | GGAGAC | 120 |
| AGA | ATTC' | TCC | GATT | GAGA | AC G | ATGA | GAGA | c cg | GAGA | GCAC | GAG | CTCC | ACA . | AACG(| CTATAG | 180 |
| ACG | CTGA | GTA | TCTG | GCGT | TG C | GTTT | GGCG | g ag | TAAA | TGGA | GAG | GAAG | AAA | TCGG | AGAGGT | 240 |
| CCA | CTTA' | TCT | AATC | GCTG | CT A | TGTT | GTCG | A GC | TTTG | GTAT | CAC | TTCT | ATG | GCTG | TTATGG | 300 |
| CTG | TTTA | CTA | CAGA | ттст | ст т | GGCA | AATG | g Ag | GGAG | GTGA | GAT | CTCA | ATG | TTGG. | AAATGT | 360 |
| TTG | GTAC. | АТТ | TGCT | CTCT | CT G | TTGG | TGCT | G CT | GTTG | GTAT | GGA | ATTC | TGG | GCAA | GATGGG | 420 |

CTCATAGAGC TCTGTGGCAC GCTTCTCTAT GGAATATGCA TGAGTCACAT CACAAACCAA

| WO 99/63055 | | | | PCT/US99/12121 |
|---|----------------------|-----------------------|------------------------|----------------|
| GAGAAGGACC GTTTGAGC | TA AACGATGTTT | TTGCTATAGT GAA | CGCTGGT CCAGCGAT | rg 540 |
| GTCTCCTCTC TTATGGAT | TC TTCAATAAAG | GACTCGTTCC TGG | TCTCTGC TTTGGCGCC | CG 600 |
| GGTTAGGCAT AACGGTGT | TT GGAATCGCCT | ACATGTTTGT CCA | CGATGGT CTCGTGCA | CA 660 |
| AGCGTTTCCC TGTAGGTC | CC ATCGCCGACG | TCCCTTACCT CCG | AAAGGTC GCCGCCGCT | °C 720 · |
| ACCAGCTACA TCACACAG | AC AAGTTCAATG | GTGTACCATA TGG | ACTGTTT CTTGGACCO | CA 780 |
| AGGAATTGGA AGAAGTTG | GA GGAAATGAAG | AGTTAGATAA GGA | GATTAGT CGGAGAATO | CA 840 |
| AATCATACAA AAAGGCCT | CG GGCTCCGGGT | CGAGTTCGAG TTC | TTGACTT TAAACAAGI | T 900 |
| TTAAATCCCA AATTCTTT | TT TTGTCTTCTG | TCATTATGAT CAT | CTTAAGA CGGTCT | 956 |
| <210> 4 <211> 294 <212> PRT <213> Arabidopsis ti | haliana | | | |
| <400> 4 Ser Phe Ser Ser Ser 1 5 | Ser Thr Asp P | he Arg Leu Arg 10 | Leu Pro Lys Ser 15 | |
| Leu Ser Gly Phe Ser 20 | Pro Ser Leu A | arg Phe Lys Arg 25 | Phe Ser Val Cys | |
| Tyr Val Val Glu Glu 35 | Arg Arg Gln A | sn Ser Pro Ile | Glu Asn Asp Glu 45 | |
| Arg Pro Glu Ser Thr 50 | Ser Ser Thr A 55 | sn Ala Ile Asp 60 | Ala Glu Tyr Leu | |
| Ala Leu Arg Leu Ala 65 | Glu Lys Leu G 70 | lu Arg Lys Lys 75 | Ser Glu Arg Ser 80 | |
| Thr Tyr Leu Ile Ala 85 | Ala Met Leu S | er Ser Phe Gly 90 | Ile Thr Ser Met 95 | |
| Ala Val Met Ala Val 100 | | he Ser Trp Gln 05 | Met Glu Gly Gly 110 | |
| Glu Ile Ser Met Leu 115 | Glu Met Phe G 120 | ly Thr Phe Ala | Leu Ser Val Gly 125 | |
| Ala Ala Val Gly Met 130 | Glu Phe Trp A 135 | la Arg Trp Ala 140 | His Arg Ala Leu | |
| Trp His Ala Ser Leu 145 | Trp Met Asn H 150 | is Glu Ser His 155 | His Lys Pro Arg 160 | |

Glu Gly Pro Phe Glu Leu Asn Asp Val Phe Ala Ile Val Asn Ala Gly

Pro Ala Ile Gly Leu Leu Ser Tyr Gly Phe Phe Asn Lys Gly Leu Val 180 185 190

Pro Gly Leu Cys Phe Gly Ala Gly Leu Gly Ile Thr Val Phe Gly Ile 195 200



Ala Tyr Met Phe Val His Asp Gly Leu Val His Lys Arg Phe Pro Val 210 215 220

Gly Pro Ile Ala Asp Val Pro Tyr Leu Arg Lys Val Ala Ala Ala His 225 230 235 240

Gln Leu His His Thr Asp Lys Phe Asn Gly Val Pro Tyr Gly Leu Phe 245 250 255

Leu Gly Pro Lys Glu Leu Glu Glu Val Gly Gly Asn Glu Glu Leu Asp 260 265 270

Lys Glu Ile Ser Arg Arg Ile Lys Ser Tyr Lys Lys Ala Ser Gly Ser 275 280 285

Gly Ser Ser Ser Ser Ser 290

<210> 5

<211> 162

<212> PRT

<213> Alicalgenes sp.

<400> 5

Met Thr Gln Phe Leu Ile Val Val Ala Thr Val Leu Val Met Glu Leu 1 5 10 15

Thr Ala Tyr Ser Val His Arg Trp Ile Met His Gly Pro Leu Gly Trp 20 25 30

Gly Trp His Lys Ser His His Glu Glu His Asp His Ala Leu Glu Lys 35 40 45

Asn Asp Leu Tyr Gly Val Val Phe Ala Val Leu Ala Thr Ile Leu Phe 50 55 60

Thr Val Gly Ala Tyr Trp Trp Pro Val Leu Trp Trp Ile Ala Leu Gly 65 70 75 80

Met Thr Val Tyr Gly Leu Ile Tyr Phe Ile Leu His Asp Gly Leu Val 85 90 95

His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Arg Gly Tyr Phe Arg 100 105 110

Arg Leu Tyr Gln Ala His Arg Leu His His Ala Val Glu Gly Arg Asp 115 120 125

His Cys Val Ser Phe Gly Phe Ile Tyr Ala Pro Pro Val Asp Lys Leu 130 135 140

Lys Gln Asp Leu Lys Arg Ser Gly Val Leu Arg Pro Gln Asp Glu Arg 145 150 155 160

Pro Ser

<210> 6

<211> 175

<212> PRT

<213> Erwinia herbicola

<400> 6 Met Leu Asn Ser Leu Ile Val Ile Leu Ser Val Ile Ala Met Glu Gly Ile Ala Ala Phe Thr His Arg Tyr Ile Met His Gly Trp Gly Trp Arg Trp His Glu Ser His His Thr Pro Arg Lys Gly Val Phe Glu Leu Asn Asp Leu Phe Ala Val Val Phe Ala Gly Val Ala Ile Ala Leu Ile Ala Val Gly Thr Ala Gly Val Trp Pro Leu Gln Trp Ile Gly Cys Gly Met Thr Val Tyr Gly Leu Leu Tyr Phe Leu Val His Asp Gly Leu Val His Gln Arg Trp Pro Phe His Trp Ile Pro Arg Arg Gly Tyr Leu Lys Arg 105 Leu Tyr Val Ala His Arg Leu His His Ala Val Arg Gly Arg Glu Gly 120 Cys Val Ser Phe Gly Phe Ile Tyr Ala Arg Lys Pro Ala Asp Leu Gln 135 Ala Ile Leu Arg Glu Arg His Gly Arg Pro Pro Lys Arg Asp Ala Ala 155 Lys Asp Arg Pro Asp Ala Ala Ser Pro Ser Ser Ser Pro Glu <210> 7 <211> 175 <212> PRT <213> Erwinia uredovora <400> 7 Met Leu Trp Ile Trp Asn Ala Leu Ile Val Phe Val Thr Val Ile Gly Met Glu Val Ile Ala Ala Leu Ala His Lys Tyr Ile Met His Gly Trp 25 Gly Trp Gly Trp His Leu Şer His His Glu Pro Arg Lys Gly Ala Phe Glu Val Asn Asp Leu Tyr Ala Val Val Phe Ala Ala Leu Ser Ile Leu

Leu Ile Tyr Leu Gly Ser Thr Gly Met Trp Pro Leu Gln Trp Ile Gly

Ala Gly Met Thr Ala Tyr Gly Leu Leu Tyr Phe Met Val His Asp Gly

Leu Val His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr

105

100

PCT/US99/12121

60

120

WO 99/63055



Leu Lys Arg Leu Tyr Met Ala His Arg Met His His Ala Val Arg Gly
115 120 125

Lys Glu Gly Cys Val Ser Phe Gly Phe Leu Tyr Ala Pro Pro Leu Ser 130 140

Lys Leu Gln Ala Thr Leu Arg Glu Arg His Gly Ala Arg Ala Gly Ala 145 150 155 160

Ala Arg Asp Ala Gln Gly Gly Glu Asp Glu Pro Ala Ser Gly Lys 165 170 175

<210> 8

<211> 162

<212> PRT

<213> Agrobacterium aurianticum

<400> 8

Met Thr Asn Phe Leu Ile Val Val Ala Thr Val Leu Val Met Glu Leu 1 5 10 15

Thr Ala Tyr Ser Val His Arg Trp Ile Met His Gly Pro Leu Gly Trp
20 25 30

Gly Trp His Lys Ser His His Glu Glu His Asp His Ala Leu Glu Lys 35 40 45

Asn Asp Leu Tyr Gly Leu Val Phe Ala Val Ile Ala Thr Val Leu Phe 50 55 60

Thr Val Gly Trp Ile Trp Ala Pro Val Leu Trp Trp Ile Ala Leu Gly 65 70 75 80

Met Thr Val Tyr Gly Leu Ile Tyr Phe Val Leu His Asp Gly Leu Val 85 90 95

His Trp Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr Ala Arg
100 105 110

Arg Leu Tyr Gln Ala His Arg Leu His His Ala Val Glu Gly Arg Asp 115 120 125

His Cys Val Ser Phe Gly Phe Ile Tyr Ala Pro Pro Val Asp Lys Leu 130 135 140

Lys Gln Asp Leu Lys Met Ser Gly Val Leu Arg Ala Glu Ala Gln Glu 145 150 155 160

Arg Thr

<210> 9

<211> 954

<212> DNA

<213> Arabidopsis thaliana

<400> 9

CCACGGGTCC GCCTCCCCGT TTTTTTCCGA TCCGATCTCC GGTGCCGAGG ACTCAGCTGT

TTGTTCGCGC TTTCTCAGCC GTCACCATGA CCGATTCTAA CGATGCTGGA ATGGATGCTG

| WO 99/63055 | | PCT/US99/12121 |
|-------------|---|----------------|
| | - | |

| TTCAGAGACG | ACTCATGTTT | GAAGACGAAT | GCATTCTCGT | TGATGAAAAT | AATCGTGTGG | 180 |
|--|------------|---------------------|----------------|-------------|---------------|-----|
| TGGGACATGA | CACTAAGTAT | AACTGTCATC | TGATGGAAAA | GATTGAAGCT | GAGAATTTAC | 240 |
| TTCACAGAGC | TTTCAGTGTG | TTTTTATTCA | ACTCCAAGTA | TGAGTTGCTT | CTCCAGCAAC | 300 |
| GGTCAAAAAC | AAAGGTTACT | TTCCCACTTG | TGTGGACAAA | CACTTGTTGC | AGCCATCCTC | 360 |
| TTTACCGTGA | ATCCGAGCTT | ATTGAAGAGA | ATGTGCTTGG | TGTAAGAAAT | GCCGCACAAA | 420 |
| GGAAGCTTTT | CGATGAGCTC | GGTATTGTAG | CAGAAGATGT | ACCAGTCGAT | GAGTTCACTC | 480 |
| CCTTGGGACG | CATGCTTTAC | AAGGCACCTT | CTGATGGGAA | ATGGGGAGAG | CACGAAGTTG | 540 |
| ACTATCTACT | CTTCATCGTG | CGGGÄTGTGA | AGCTTCAACC | AAACCCAGAT | GAAGTGGCTG | 600 |
| AGATCAAGTA | CGTGAGCAGG | GAAGAGCTTA | AGGAGCTGGT | GAAGAAAGCA | GATGCTGGCG | 660 |
| ATGAAGCTGT | GAAACTATCT | CCATGGTTCA | GATTGGTGGT | GGATAATTTC | TTGATGAAGT | 720 |
| GGTGGGATCA | TGTTGAGAAA | GGAACTATCA | CTGAAGCTGC | AGACATGAAA | ACCATTCACA | 780 |
| AGCTCTGAAC | TTTCCATAAG | TTTTGGATCT | TCCCCTTCCC | ATAATAAAT | TAAGAGATGA | 840 |
| GACTTTTATT | GATTACAGAC | AAAACTGGCA | ACAAAATCTA | TTCCTAGGAT | TTTTTTTGC | 900 |
| TTTTTTTTTA | CTTTTGATTC | ATCTCTAGTT | TAGTTTTCAT | CTTAAAAAAA | AAAA | 954 |
| <210> 10 <211> 996 <212> DNA <213> Arabidopsis thaliana | | | | | | |
| <400> 10 | TCTTTCTTCT | ## ##### ### | TCCCA TITCA TI | mcccoman en | momomoo e e e | |
| | TGTTTCTTCT | | | | | 60 |
| | TTTTTCTTCT | | | | | 120 |
| | GAATTTTCGT | | | | | 180 |
| | TGTTCAGAGA | | | | | 240 |
| | TGTGGGGCAT | | | | | 300 |
| | | | | | TATGAGTTGC | 360 |
| | AAGGTCAAAC | | | | | 420 |
| | | | | | GGTGTGAGGA | 480 |
| ATGCTGCACA | AAGAAAGCTT | CTCGATGAGC | TTGGTATTGT | AGCTGAAGAT | GTACCAGTCG | 540 |

600

660

720

780

ATGAGTTCAC TCCCTTGGGA CGTATGCTGT ACAAGGCTCC TTCTGATGGC AAATGGGGAG

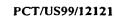
AGCATGAACT TGATTACTTG CTCTTCATCG TGCGAGACGT GAAGGTTCAA CCAAACCCAG

ATGAAGTAGC TGAGATCAAG TATGTGAGCC GGGAAGAGCT GAAGGAGCTG GTGAAGAAAG

CAGATGCAGG TGAGGAAGGT TTGAAACTGT CACCATGGTT CAGATTGGTG GTGGACAATT

WO 99/63055

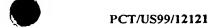




| | _ | | | | | |
|--|--------------|------------|------------|------------|------------|------|
| TCTTGATGAA | GTGGTGGGAT | CATGTTGAGA | AAGGAACTTT | GGTTGAAGCT | ATAGACATGA | 840 |
| AAACCATCCA | CAAACTCTGA | ACATCTTTTT | TTAAAGTTTT | TAAATCAATC | AACTTTCTCT | 900 |
| TCATCATTTT | TATCTTTTCG | ATGATAATAA | TTTGGGATAT | GTGAGACACT | TACAAAACTT | 960 |
| CCAAGCACCT | CAGGCAATAA | TAAAGTTTGC | GGCCGC | | | 996 |
| <210> 11 <211> 1165 <212> DNA <213> Haema | atococcus pl | uvialis | | | \ | |
| <400> 11 CTCGGTAGCT | GGCCACAATC | GCTATTTGGA | ACCTGGCCCG | GCGGCAGTCC | GATGCCGCGA | 60 |
| TGCTTCGTTC | GTTGCTCAGA | GGCCTCACGC | ATATCCCCCG | CGTGAACTCC | GCCCAGCAGC | 120 |
| CCAGCTGTGC | ACACGCGCGA | CTCCAGTTTA | AGCTCAGGAG | CATGCAGATG | ACGCTCATGC | 180 |
| AGCCCAGCAT | CTCAGCCAAT | CTGTCGCGCG | CCGAGGACCG | CACAGACCAC | ATGAGGGGTG | 240 |
| CAAGCACCTG | GGCAGGCGGG | CAGTCGCAGG | ATGAGCTGAT | GCTGAAGGAC | GAGTGCATCT | 300 |
| TGGTGGATGT | TGAGGACAAC | ATCACAGGCC | ATGCCAGCAA | GCTGGAGTGT | CACAAGTTCC | 360 |
| TACCACATCA | GCCTGCAGGC | CTGCTGCACC | GGGCCTTCTC | TGTGTTCCTG | TTTGACGATC | 420 |
| AGGGGCGACT | GCTGCTGCAA | CAGCGTGCAC | GCTCAAAAAT | CACCTTCCCA | AGTGTGTGGA | 480 |
| CGAACACCTG | CTGCAGCCAC | CCTTTACATG | GGCAGACCCC | AGATGAGGTG | GACCAACTAA | 540 |
| GCCAGGTGGC | CGACGGAACA | GTACCTGGCĠ | CAAAGGCTGC | TGCCATCCGC | AAGTTGGAGC | 600 |
| ACGAGCTGGG | GATACCAGCG | CACCAGCTGC | CGGCAAGCGC | GTTTCGCTTC | CTCACGCGTT | 660 |
| TGCACTACTG | TGCCGCGGAC | GTGCAGCCAG | CTGCGACACA | ATCAGCGCTC | TGGGGCGAGC | 720 |
| ACGAAATGGA | CTACATCTTG | TTCATCCGGG | CCAACGTCAC | CTTGGCGCCC | AACCCTGACG | 780 |
| AGGTGGACGA | AGTCAGGTAC | GTGACGCAAG | AGGAGCTGCG | GCAGATGATG | CAGCCGGACA | 840 |
| ACGGGCTGCA | ATGGTCGCCG | TGGTTTCGCA | TCATCGCCGC | GCGCTTCCTT | GAGCGTTGGT | 900 |
| GGGCTGACCT | GGACGCGGCC | CTAAACACTG | ACAAACACGA | GGATTGGGGA | ACGGTGCATC | 960 |
| ACATCAACGA | AGCGTGAAAG | CAGAAGCTGC | AGGATGTGAA | GACACGTCAT | GGGGTGGAAT | 1020 |
| TGCGTACTTG | GCAGCTTCGT | ATCTCCTTTT | TCTGAGACTG | AACCTGCAGT | CAGGTCCCAC | 1080 |
| AAGGTCAGGT | AAAATGGCTC | GATAAAATGT | ACCGTCACTT | TTTGTCGCGT | ATACTGAACT | 1140 |
| CCAAGAGGTC | AAAAAAAAA | AAAAA | | | | 1165 |
| | | | | | | |

<210> 12 <211> 1135 <212> DNA <213> Haematococcus pluvialis

| wo | 00/63055 | 1 |
|----|----------|---|
| wo | 99/63055 | |



480

| <400> 12 | | | | | ين جو | |
|--|------------|------------|------------|------------|------------|------|
| | GGCCACAATC | GCTATTTGGA | ACCTGGCCCG | GCGGCAGTCC | GATGCCGCGA | 60 |
| TGCTTCGTTC | GTTGCTCAGA | GGCCTCACGC | ATATCCCGCG | CGTGAACTCC | GCCCAGCAGC | 120 |
| CCAGCTGTGC | ACACGCGCGA | CTCCAGTTTA | AGCTCAGGAG | CATGCAGCTG | CTTTCCGAGG | 180 |
| ACCGCACAGA | CCACATGAGG | GGTGCAAGCA | CCTGGGCAGG | CGGGCAGTCG | CAGGATGAGC | 240 |
| TGATGCTGAA | GGACGAGTGC | ATCTTGGTAG | ATGTTGAGGA | CAACATCACA | GGCCATGCCA | 300 |
| GCAAGCTGGA | GTGTCACAAG | TTCCTACCAC | ATCAGCCTGC | AGGCCTGCTG | CACCGGGCCT | 360 |
| TCTCTGTGTT | CCTGTTTGAC | GATCAGGGGC | GACTGCTGCT | GCAACAGCGT | GCACGCTCAA | 420 |
| AAATCACCTT | CCCAAGTGTG | TGGACGAACA | CCTGCTGCAG | CCACCCTTTA | CATGGGCAGA | 480 |
| CCCCAGATGA | GGTGGACCAA | CTAAGCCAGG | TGGCCGACGG | AACAGTACCT | GGCGCAAAGG | 540 |
| CTGCTGCCAT | CCGCAAGTTG | GAGCACGAGC | TGGGGATACC | AGCGCACCAG | CTGCCGGCAA | 600 |
| GCGCGTTTCG | CTTCCTCACG | CGTTTGCACT | ACTGTGCCGC | GGACGTGCAG | CCAGCTGCGA | 660 |
| CACAATCAGC | GCTCTGGGGC | GAGCACGAAA | TGGACTACAT | CTTGTTCATC | CGGGCCAACG | 720 |
| TCACCTTGGC | GCCCAACCCT | GACGAGGTGG | ACGAAGTCAG | GTACGTGACG | CAAGAGGAGC | 780 |
| TGCGGCAGAT | GATGCAGCCG | GACAACGGGC | TTCAATGGTC | GCCGTGGTTT | CGCATCATCG | 840 |
| CCGCGCGCTT | CCTTGAGCGT | TGGTGGGCTG | ACCTGGACGC | GGCCCTAAAC | ACTGACAAAC | 900 |
| ACGAGGATTG | GGGAACGGTG | CATCACATCA | ACGAAGCGTG | AAGGCAGAAG | CTGCAGGATG | 960 |
| TGAAGACACG | TCATGGGGTG | GAATTGCGTA | CTTGGCAGCT | TCGTATCTCC | TTTTTCTGAG | 1020 |
| ACTGAACCTG | CAGAGCTAGA | GTCAATGGTG | CATCATATTC | ATCGTCTCTC | TTTTGTTTTA | 1080 |
| GACTAATCTG | TAGCTAGAGT | CACTGATGAA | TCCTTTACAA | CTTTCAAAAA | AAAAA ' | 1135 |
| <210> 13 <211> 960 <212> DNA <213> Tagetes erecta | | | | | | |
| <400> 13 CCAAAAACAA | CTCAAATCTC | CTCCGTCGCT | CTTACTCCGC | CATGGGTGAC | GACTCCGGCA | 60 |
| TGGATGCTGT | TCAGCGACGT | CTCATGTTTG | ACGATGAATG | CATTTTGGTG | GATGAGTGTG | 120 |
| ACAATGTGGT | GGGACATGAT | ACCAAATACA | ATTGTCACTT | GATGGAGAAG | ATTGAAACAG | 180 |
| GTAAAATGCT | GCACAGAGCA | TTCAGCGTTT | TTCTATTCAA | TTCAAAATAC | GAGTTACTTC | 240 |
| TTCAGCAACG | GTCTGCAACC | AAGGTGACAT | TTCCTTTAGT | ATGGACCAAC | ACCTGTTGCA | 300 |
| GCCATCCACT | CTACAGAGAA | TCCGAGCTTG | TTCCCGAAAC | GCCTGAGAGA | ATGCTGCACA | 360 |
| GAGGANNNN | иииииииии | ииииииииии | ииииииииии | ииииииииии | ииииииииии | 420 |

имининий имининини имининини имининини имининини имининини

PCT/US99/12121

WO 99/63055

| ииииииииии | ииииииииии | иииииииии | иииииииии | ииииииииии | иииииииии | 540 |
|------------|------------|------------|------------|------------|------------|-----|
| ииииииииии | иииииииии | иииииииии | ииийииииии | иииииииии | ииииииииии | 600 |
| имимимими | ииииииииии | ииииииииии | иииииииии | иииииииии | имимимими | 660 |
| ииииииииии | иииииииии | TCATGTGCAA | AAGGGTACAC | TCACTGAATG | CAATTTGATA | 720 |
| TGAAAACCAT | ACACAAGCTG | ATATAGAAAC | ACACCCTCAA | CCGAAAAGCA | AGCCTAATAA | 780 |
| TTCGGGTTGG | GTCGGGTCTA | CCATCAATTG | TTTTTTTTTT | TTAACAACTT | TTAATCTCTA | 840 |
| TTTGAGCATG | TTGATTCTTG | TCTTTTGTGT | GTAAGATTTT | GGGTTTCGTT | TCAGTTGTAA | 900 |
| TAATGAACCA | TTGATGGTTT | GCAATTTCAA | GTTCCTATCG | ACATGTAGTG | АТСТАААААА | 960 |
| | | | | | | |

<210> 14

<211> 305

<212> PRT

<213> Haematococcus pluvialis

<400> 14

Met Leu Arg Ser Leu Leu Arg Gly Leu Thr His Ile Pro Arg Val Asn 1 5 10 15

Ser Ala Gln Gln Pro Ser Cys Ala His Ala Arg Leu Gln Phe Lys Leu 20 25 30

Arg Ser Met Gln Met Thr Leu Met Gln Pro Ser Ile Ser Ala Asn Leu 35 40 45

Ser Arg Ala Glu Asp Arg Thr Asp His Met Arg Gly Ala Ser Thr Trp 50 55 60

Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys Asp Glu Cys Ile 65 70 75 80

Leu Val Asp Val Glu Asp Asn Ile Thr Gly His Ala Ser Lys Leu Glu 85 90 95

Cys His Lys Phe Leu Pro His Gln Pro Ala Gly Leu Leu His Arg Ala 100 105 110

Phe Ser Val Phe Leu Phe Asp Asp Gln Gly Arg Leu Leu Gln Gln 115 120 125

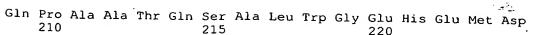
Arg Ala Arg Ser Lys Ile Thr Phe Pro Ser Val Trp Thr Asn Thr Cys 130 135 140

Cys Ser His Pro Leu His Gly Gln Thr Pro Asp Glu Val Asp Gln Leu 145 150 155 160

Ser Gln Val Ala Asp Gly Thr Val Pro Gly Ala Lys Ala Ala Ile 165 170 175

Arg Lys Leu Glu His Glu Leu Gly Ile Pro Ala His Gln Leu Pro Ala 180 185 190

Ser Ala Phe Arg Phe Leu Thr Arg Leu His Tyr Cys Ala Ala Asp Val 195 200 205



- Tyr Ile Leu Phe Ile Arg Ala Asn Val Thr Leu Ala Pro Asn Pro Asp 225 230 235 240
- Glu Val Asp Glu Val Arg Tyr Val Thr Gln Glu Glu Leu Arg Gln Met 245 250 255
- Met Gln Pro Asp Asn Gly Leu Gln Trp Ser Pro Trp Phe Arg Ile Ile 260 265 270
- Ala Ala Arg Phe Leu Glu Arg Trp Trp Ala Asp Leu Asp Ala Ala Leu 275 280 285
- Asn Thr Asp Lys His Glu Asp Trp Gly Thr Val His His Ile Asn Glu 290 295 300

Ala 305

- <210> 15
- <211> 293
- <212> PRT
- <213> Haematococcus pluvialis

<400> 15

- Met Leu Arg Ser Leu Leu Arg Gly Leu Thr His Ile Pro Arg Val Asn
 1 5 10 15
- Ser Ala Gln Gln Pro Ser Cys Ala His Ala Arg Leu Gln Phe Lys Leu 20 25 30
- Arg Ser Met Gln Leu Leu Ser Glu Asp Arg Thr Asp His Met Arg Gly
 35 40 45
- Ala Ser Thr Trp Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys 50 55 60
- Asp Glu Cys Ile Leu Val Asp Val Glu Asp Asn Ile Thr Gly His Ala 65 70 75
- Ser Lys Leu Glu Cys His Lys Phe Leu Pro His Gln Pro Ala Gly Leu 85 90 95
- Leu His Arg Ala Phe Ser Val Phe Leu Phe Asp Asp Gln Gly Arg Leu 100 105 110
- Leu Leu Gln Gln Arg Ala Arg Ser Lys Ile Thr Phe Pro Ser Val Trp 115 120 125
- Thr Asn Thr Cys Cys Ser His Pro Leu His Gly Gln Thr Pro Asp Glu 130 135 140
- Val Asp Gln Leu Ser Gln Val Ala Asp Gly Thr Val Pro Gly Ala Lys
 145 150 155 160
- Ala Ala Ala Ile Arg Lys Leu Glu His Glu Leu Gly Ile Pro Ala His 165 170 175
- Gln Leu Pro Ala Ser Ala Phe Arg Phe Leu Thr Arg Leu His Tyr Cys



180 185 1

Ala Ala Asp Val Gln Pro Ala Ala Thr Gln Ser Ala Leu Trp Gly Glu 195 200 205

His Glu Met Asp Tyr Ile Leu Phe Ile Arg Ala Asn Val Thr Leu Ala 210 215 220

Pro Asn Pro Asp Glu Val Asp Glu Val Arg Tyr Val Thr Gln Glu Glu 225 230 235 240

Leu Arg Gln Met Met Gln Pro Asp Asn Gly Leu Gln Trp Ser Pro Trp . 245 250 255

Phe Arg Ile Ile Ala Ala Arg Phe Leu Glu Arg Trp Trp Ala Asp Leu 260 265 270

Asp Ala Ala Leu Asn Thr Asp Lys His Glu Asp Trp Gly Thr Val His 275 280 285

His Ile Asn Glu Ala 290

<210> 16

<211> 284

<212> PRT

<213> Arabidopsis thaliana

<400> 16

BNSDOCID: <WO 996305541 I

Met Ser Val Ser Ser Leu Phe Asn Leu Pro Leu Ile Arg Leu Arg Ser 1 5 10 15

Leu Ala Leu Ser Ser Ser Phe Ser Ser Phe Arg Phe Ala His Arg Pro 20 25 30

Leu Ser Ser Ile Ser Pro Arg Lys Leu Pro Asn Phe Arg Ala Phe Ser 35 40 45

Gly Thr Ala Met Thr Asp Thr Lys Asp Ala Gly Met Asp Ala Val Gln 50 55 60

Arg Arg Leu Met Phe Glu Asp Glu Cys Ile Leu Val Asp Glu Thr Asp 65 70 75 80

Arg Val Val Gly His Val Ser Lys Tyr Asn Cys His Leu Met Glu Asn 85 90 95

Ile Glu Ala Lys Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe 100 105 110

Asn Ser Lys Tyr Glu Leu Leu Gln Gln Arg Ser Asn Thr Lys Val 115 120 125

Thr Phe Pro Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr 130 135 140

Arg Glu Ser Glu Leu Ile Gln Asp Asn Ala Leu Gly Val Arg Asn Ala 145 150 155 160

Ala Gln Arg Lys Leu Leu Asp Glu Leu Gly Ile Val Ala Glu Asp Val 165 170 175



Pro Val Asp Glu Phe Thr Pro Leu Gly Arg Met Leu Tyr Lys Ala Pro 180 185 190

Ser Asp Gly Lys Trp Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Ile 195 200 205

Val Arg Asp Val Lys Val Gln Pro Asn Pro Asp Glu Val Ala Glu Ile 210 215 220

Lys Tyr Val Ser Arg Glu Glu Leu Lys Glu Leu Val Lys Lys Ala Asp 225 230 235 240

Ala Gly Glu Gly Leu Lys Leu Ser Pro Trp Phe Arg Leu Val Val 245 250 255

Asp Asn Phe Leu Met Lys Trp Trp Asp His Val Glu Lys Gly Thr Leu 260 265 270

Val Glu Ala Ile Asp Met Lys Thr Ile His Lys Leu 275 280

<210> 17

<211> 287

<212> PRT

<213> Clarkia breweri

<400> 17

Met Ser Ser Ser Met Leu Asn Phe Thr Ala Ser Arg Ile Val Ser Leu 1 5 10 15

Pro Leu Leu Ser Ser Pro Pro Ser Arg Val His Leu Pro Leu Cys Phe 20 25 30

Phe Ser Pro Ile Ser Leu Thr Gln Arg Phe Ser Ala Lys Leu Thr Phe 35 40 45

Ser Ser Gln Ala Thr Thr Met Gly Glu Val Val Asp Ala Gly Met Asp 50 55 60

Ala Val Gln Arg Arg Leu Met Phe Glu Asp Glu Cys Ile Leu Val Asp 65 70 75 80

Glu Asn Asp Lys Val Val Gly His Glu Ser Lys Tyr Asn Cys His Leu 85 90 95

Met Glu Lys Ile Glu Ser Glu Asn Leu Leu His Arg Ala Phe Ser Val 100 105 110

Phe Leu Phe Asn Ser Lys Tyr Glu Leu Leu Gln Gln Arg Ser Ala 115 120 125

Thr Lys Val Thr Phe Pro Leu Val Trp Thr Asn Thr Cys Cys Ser His 130 135 140

Pro Leu Tyr Arg Glu Ser Glu Leu Ile Asp Glu Asn Cys Leu Gly Val 145 150 155 160

Arg Asn Ala Ala Gln Arg Lys Leu Leu Asp Glu Leu Gly Ile Pro Ala 165 170 175

Glu Asp Leu Pro Val Asp Gln Phe Ile Pro Leu Ser Arg Ile Leu Tyr

180

185

190

Lys Ala Pro Ser Asp Gly Lys Trp Gly Glu His Glu Leu Asp Tyr Leu 195 200 205

Leu Phe Ile Ile Arg Asp Val Asn Leu Asp Pro Asn Pro Asp Glu Val 210 215 220

Ala Glu Val Lys Tyr Met Asn Arg Asp Asp Leu Lys Glu Leu Leu Arg 225 230 235 240

Lys Ala Asp Ala Glu Glu Glu Gly Val Lys Leu Ser Pro Trp Phe Arg 245 250 255

Leu Val Val Asp Asn Phe Leu Phe Lys Trp Trp Asp His Val Glu Lys 260 265 270

Gly Ser Leu Lys Asp Ala Ala Asp Met Lys Thr Ile His Lys Leu 275 280 285

<210> 18

<211> 261

<212> PRT

<213> Arabidopsis thaliana

<400> 18

Thr Gly Pro Pro Pro Arg Phe Pro Ile Arg Ser Pro Val Pro Arg

1 5 10 15

Thr Gln Leu Phe Val Arg Ala Phe Ser Ala Val Thr Met Thr Asp Ser 20 25 30

Asn Asp Ala Gly Met Asp Ala Val Gln Arg Arg Leu Met Phe Glu Asp $35 \hspace{1cm} 40 \hspace{1cm} 45$

Glu Cys Ile Leu Val Asp Glu Asn Asn Arg Val Val Gly His Asp Thr
50 55 60

Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala Glu Asn Leu Leu 65 70 75 80

His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu Leu Leu 85 90 95

Leu Gln Gln Arg Ser Lys Thr Lys Val Thr Phe Pro Leu Val Trp Thr
100 105 110

Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser Glu Leu Ile Glu 115 120 125

Glu Asn Val Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu Phe Asp 130 135 140

Glu Leu Gly Ile Val Ala Glu Asp Val Pro Val Asp Glu Phe Thr Pro 145 150 155 160

Leu Gly Arg Met Leu Tyr Lys Ala Pro Ser Asp Gly Lys Trp Gly Glu
165 170 175

His Glu Val Asp Tyr Leu Leu Phe Ile Val Arg Asp Val Lys Leu Gln 180 185 190 Pro Asn Pro Asp Glu Val Ala Glu Ile Lys Tyr Val Ser Arg Glu Glu 195 200 205

Leu Lys Glu Leu Val Lys Lys Ala Asp Ala Gly Asp Glu Ala Val Lys 210 220

Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe Leu Met Lys Trp 225 230 235 240

Trp Asp His Val Glu Lys Gly Thr Ile Thr Glu Ala Ala Asp Met Lys 245 250 255

Thr Ile His Lys Leu 260

<210> 19

<211> 288

<212> PRT

<213> Saccharomyces cerevisiae

<400> 19

Met Thr Ala Asp Asn Asn Ser Met Pro His Gly Ala Val Ser Ser Tyr

1 10 15

Ala Lys Leu Val Gln Asn Gln Thr Pro Glu Asp Ile Leu Glu Glu Phe 20 25 30

Pro Glu Ile Ile Pro Leu Gln Gln Arg Pro Asn Thr Arg Ser Ser Glu 35 40 45

Thr Ser Asn Asp Glu Ser Gly Glu Thr Cys Phe Ser Gly His Asp Glu
50 55 60

Glu Gln Ile Lys Leu Met Asn Glu Asn Cys Ile Val Leu Asp Trp Asp 65 70 75 80

Asp Asn Ala Ile Gly Ala Gly Thr Lys Lys Val Cys His Leu Met Glu 85 90 95

Asn Ile Glu Lys Gly Leu Leu His Arg Ala Phe Ser Val Phe Ile Phe 100 105 110

Asn Glu Gln Gly Glu Leu Leu Gln Gln Arg Ala Thr Glu Lys Ile 115 120 125

Thr Phe Pro Asp Leu Trp Thr Asn Thr Cys Cys Ser His Pro Leu Cys 130 135 140

Ile Asp Asp Glu Leu Gly Leu Lys Gly Lys Leu Asp Asp Lys Ile Lys 145 150 155 160

Gly Ala Ile Thr Ala Ala Val Arg Lys Leu Asp His Glu Leu Gly Ile 165 170 175

Pro Glu Asp Glu Thr Lys Thr Arg Gly Lys Phe His Phe Leu Asn Arg 180 185 190

Ile His Tyr Met Ala Pro Ser Asn Glu Pro Trp Gly Glu His Glu Ile
195 200 205

Asp Tyr Ile Leu Phe Tyr Lys Ile Asn Ala Lys Glu Asn Leu Thr Val

210 215 220

Asn Pro Asn Val Asn Glu Val Arg Asp Phe Lys Trp Val Ser Pro Asn 225 230 235 240

Asp Leu Lys Thr Met Phe Ala Asp Pro Ser Tyr Lys Phe Thr Pro Trp 245 250 255

Phe Lys Ile Ile Cys Glu Asn Tyr Leu Phe Asn Trp Trp Glu Gln Leu 260 . 265 270

Asp Asp Leu Ser Glu Val Glu Asn Asp Arg Gln Ile His Arg Met Leu 275 280 285

<210> 20

<211> 456

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Consensus
 sequence of four plant B-cyclases '

<400> 20

Met Asp Thr Leu Leu Lys Thr Pro Asn Leu Glu Phe Leu Pro His Gly
1 5 10 15

Phe Val Lys Ser Phe Ser Lys Phe Gly Lys Cys Glu Gly Val Cys Val 20 25 30

Lys Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr Lys Lys Glu Asn 35 40 45

Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys Gly Val Val Asp
50 55 60

Leu Ala Val Val Gly Gly Pro Ala Gly Leu Ala Val Ala Gln Gln 65 70 75 80.

Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile Asp Pro Pro Lys Leu 85 90 95

Ile Trp Pro Asn Asn Tyr Gly Val Trp Val Asp Glu Phe Glu Ala Met 100 105 110

Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp Ser Gly Ala Val Tyr Ile 115 120 125

Asp Asp Thr Lys Asp Leu Arg Pro Tyr Gly Arg Val Asn Arg Lys Gln 130 135 140

Leu Lys Ser Lys Met Met Gln Lys Cys Ile Asn Gly Val Lys Phe His 145 150 155 160

Gln Ala Lys Val Ile Lys Val Ile His Glu Glu Lys Ser Met Leu Ile 165 170 175

Cys Asn Asp Gly Thr Ile Gln Ala Thr Val Val Leu Asp Ala Thr Gly 180 185 190

Phe Ser Arg Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln

- E

205

195 200

Val Ala Tyr Gly Ile Leu Ala Glu Val Glu Glu His Pro Phe Asp Lys 210 215 220

Met Val Phe Met Asp Trp Arg Asp Ser His Leu Asn Asn Glu Leu Lys 235 240

Glu Arg Asn Ser Ile Pro Thr Phe Leu Tyr Ala Met Pro Phe Ser Ser 245 250 255

Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg Pro Gly Leu 260 265 270

Arg Met Asp Asp Ile Gln Glu Arg Met Val Ala Arg Leu His Leu Gly 275 280 285

Ile Lys Val Lys Ser Ile Glu Glu Asp Glu His Cys Val Ile Pro Met 290 295 300

Gly Gly Pro Leu Pro Val Leu Pro Gln Arg Val Val Gly Ile Gly Gly 305 310 315

Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met Val Ala Arg Thr 325 330 335

Leu Ala Ala Pro Val Val Ala Asn Ala Ile Ile Tyr Leu Gly Ser 340 345 350

Glu Ser Ser Gly Glu Leu Ser Ala Glu Val Trp Lys Asp Leu Trp Pro 355 360 365

Ile Glu Arg Arg Gln Arg Glu Phe Phe Cys Phe Gly Met Asp Ile 370 375 380

Leu Leu Lys Leu Asp Leu Pro Ala Thr Arg Arg Phe Phe Asp Ala Phe 385 390 395 400

Phe Asp Leu Glu Pro Arg Tyr Trp His Gly Phe Leu Ser Ser Arg Leu 405 410 415

Phe Leu Pro Glu Leu Ile Val Phe Gly Leu Ser Leu Phe Ser His Ala 420 425 430

Ser Asn Thr Ser Arg Glu Ile Met Thr Lys Gly Thr Pro Leu Val Met 435 440 445

Ile Asn Asn Leu Leu Gln Asp Glu 450 455

<210> 21

<211> 524

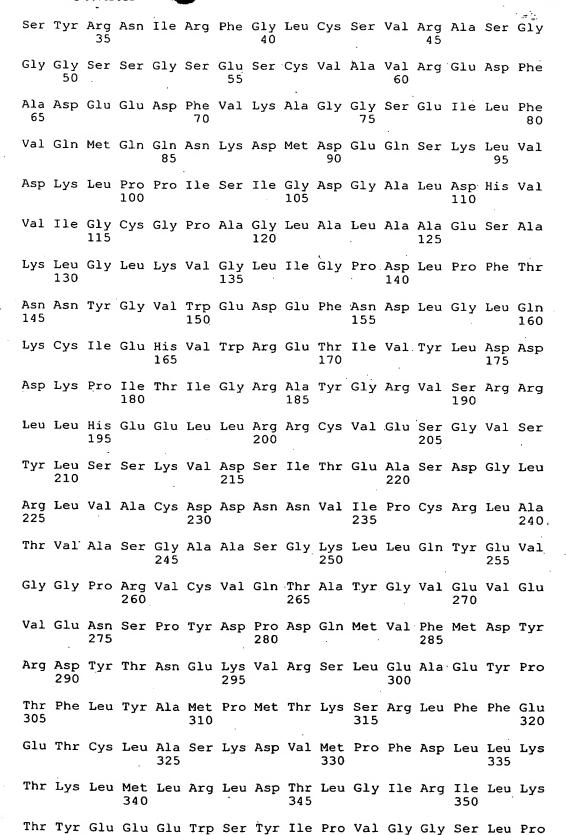
<212> PRT

<213> Arabidopsis thaliana

<400> 21

Met Glu Cys Val Gly Ala Arg Asn Phe Ala Ala Met Ala Val Ser Thr

Phe Pro Ser Trp Ser Cys Arg Arg Lys Phe Pro Val Val Lys Arg Tyr 20 25 30



WO 99/63055 PCT/US99/12121

| • | | | |
|--|------------------------|------------------------|------------|
| 355 | 360 | 365 | ario |
| Asn Thr Glu Gln Lys Asn Leu 370 375 | Ala Phe Gly Ala | Ala Ala Ser Met 380 | Val |
| His Pro Ala Thr Gly Tyr Ser 385 390 | Val Val Arg Ser 395 | Leu Ser Glu Ala | Pro 400 |
| Lys Tyr Ala Ser Val Ile Ala 405 | Glu Ile Leu Arg 410 | Glu Glu Thr Thr 415 | Lys |
| Gln Ile Asn Ser Asn Ile Ser . 420 | Arg Gln Ala Trp 425 | Asp Thr Leu Trp 430 | Pro · |
| Pro Glu Arg Lys Arg Gln Arg . 435 | Ala Phe Phe Leu 440 | Phe Gly Leu Ala | Leu |
| Ile Val Glm Phe Asp Thr Glu 450 455 | Gly Ile Arg Ser | Phe Phe Arg Thr 460 | Phe |
| Phe Arg Leu Pro Lys Trp Met 465 470 | Trp Gln Gly Phe 475 | Leu Gly Ser Thr | Leu 480 |
| Thr Ser Gly Asp Leu Val Leu 485 | Phe Ala Leu Tyr 490 | Met Phe Val Ile 495 | Ser |
| Pro Asn Asn Leu Arg Lys Gly 500 | Leu Ile Asn His 505 | Leu Ile Ser Asp 510 | Pro |
| Thr Gly Ala Thr Met Ile Lys 5 | Thr Tyr Leu Lys 520 | Val | |
| <210> 22 <211> 1898 <212> DNA <213> Adonis palaestina | | | |
| <400> 22 AAAGGAGTGT TCTATTAATG TTACTG | TCGC ATTCTTGCAA | CACTTATATT CAAAC | CTCCAT 60 |
| TTTCTTCTTT TCTCTTCAAA ACAACA | AACT AATGTGAGCA | GAGTATCTGG CTATG | GGAACT 120 |
| ACTTGGTGTT CGCAACCTCA TCTCTT | CTTG CCCTGTGTGG | ACTTTTGGAA CAAGA | AACCT 180 |
| TAGTAGTTCA AAACTAGCTT ATAACA | TACA TCGATATGGT | TCTTCTTGTA GAGTA | GATTT 240 |
| TCAAGTGAGA GCTGATGGTG GAAGCG | GGAG TAGAAGTTCT | GTTGCTTATA AAGAG | GGTTT 300 |
| TGTGGATGAA GAGGATTTTA TCAAAG | CTGG TGGTTCTGAG | CTTTTGTTTG TCCAP | ATGCA 360 |
| GCAAACAAAG TCTATGGAGA AACAGG | CCAA GCTCGCCGAT | AAGTTGCCAC CAATA | ACCTTT 420 |
| TGGAGAATCC GTGATGGACT TGGTTG | TAAT AGGTTGTGGA | CCTGCTGGTC TTTC | CTGGC 480 |
| TGCAGAAGCT GCTAAGCTAG GGTTGA | AAGT TGGCCTTATT | GGTCCTGATC TTCCT | TTTAC 540 |
| AAATAATTAT GGTGTGTGGG AAGACG | AGTT CAAAGATCTT | GGACTTGAAC GTTGT | CATCGA 600 |
| GCATGCTTGG AAGGACACCA TCGTAT | ATCT TGATAATGAT | GCTCCTGTCC TTATT | GGTCG 660 |

TGCATATGGA CGAGTTAGTC GACATTTGCT ACATGAGGAG TTGCTGAAAA GGTGTGTGGA

BNSDOCID: <WO 998305541 (>

720

WO 99/63055



| - | - | . ~ ~ | | | _ |
|-----|----|-------|------|-----|---|
| PCT | /L | JS9 | 19/1 | 121 | 2 |

| GTCAGGTGTA | TCATATCTTG | ATTCTAAAGT | GGAAAGGATC | ACTGAAGCTG | GTGATGGCCA | 780 |
|------------|------------|------------|------------|------------|------------|-------|
| TAGCCTTGTA | GTTTGTGAAA | ATGAGATCTT | TATCCCTTGC | AGGCTTGCTA | CTGTTGCATC | 840 |
| TGGAGCAGCT | TCAGGGAAAC | TTTTGGAGTA | TGAAGTAGGT | GGCCCTCGTG | TTTGTGTCCA | . 900 |
| AACCGCTTAT | GGGGTGGAGG | TTGAGGTGGA | GAACAATCCA | TACGATCCCA | ACTTAATGGT | 960 |
| ATTCATGGAC | TACAGAGACT | ATATGCAACA | GAAATTACAG | TGCTCGGAAG | AAGAATATCC | 1020 |
| AACATTTCTC | TATGTCATGC | CCATGTCGCC | AACAAGACTT | TTTTTTGAGG | AAACCTGTTT | 1080 |
| GGCCTCAAAA | GATGCCATGC | CATTCGATCT | ACTGAAGAGA | AAACTGATGT | CACGATTGAA | 1140 |
| GACTCTGGGT | ATCCAAGTTA | CAAAAGTTTA | TGAAGAGGAA | TGGTCATATA | TTCCTGTTGG | 1200 |
| TGGTTCTTTA | CCAAACACAG | AGCAAAAGAA | CCTAGCATTT | GGTGCTGCAG | CAAGCATGGT | 1260 |
| GCATCCAGCA | ACAGGCTATT | CGGTTGTACG | GTCACTGTCA | GAAGCTCCAA | AATATGCTTC | 1320 |
| TGTAATTGCA | AAGATTTTGA | AGCAAGATAA | CTCTGCGTAT | GTGGTTTCTG | GACAAAGTAG | 1380 |
| TGCAGTAAAC | ATTTCAATGC | AAGCATGGAG | CAGTCTTTGG | CCAAAGGAGC | GAAAACGTCA | 1440 |
| AAGAGCATTC | TTTCTTTTTG | GATTAGAGCT | TATTGTGCAG | CTAGATATTG | AAGCAACCAG | 1500 |
| AACATTCTTT | AGAACCTTCT | TCCGCTTGCC | AACTTGGATG | TGGTGGGGTT | TCCTTGGGTC | 1560 |
| TTCACTATCA | TCTTTCGATC | TCGTCTTGTT | TTCCATGTAC | ATGTTTGTTT | TGGCGCCAAA | 1620 |
| CAGCATGAGG | ATGTCACTTG | TGAGACATTT | GCTTTCAGAT | CCTTCTGGTG | CAGTTATGGT | 1680 |
| AAGAGCTTAC | CTCGAAAGGT | AGTCTCATCT | ATTATTAAAC | TCTAGTGTTT | CACCAAATAA | 1740 |
| ATGAGGATCC | TTCGAATGTG | TATATGATCA | TCTCTATGTA | TATCCTGTAC | TCTAATCTCA | 1800 |
| TAAAGTAAAT | GCCGGGTTTG | ATATTGTTGT | GTCAAACCGG | CCAATGATAT | AAAGTAAATT | 1860 |
| TATTGATACA | AAAGTAGTTT | TTTTCCTTAA | AAAAAAA | | • | 1898 |

<210> 23

<211> 529

<212> PRT

<213> Adonis palaestina

<400> 23

BNSDOCID: <WO 996305541 L >

Met Glu Leu Leu Gly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp
1 5 10 15

Thr Phe Gly Thr Arg Asn Leu Ser Ser Ser Lys Leu Ala Tyr Asn Ile 20 25 30

His Arg Tyr Gly Ser Ser Cys Arg Val Asp Phe Gln Val Arg Ala Asp 35 40 . 45

Gly Gly Ser Gly Ser Arg Ser Ser Val Ala Tyr Lys Glu Gly Phe Val 50 55 60

Asp Glu Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu Leu Leu Phe Val 65 70 75 80



| Gln | Met | Gln | Gln | Thr 85 | Lys | Ser | Met | Glu | Lys 90 | Gln | Ala | Lys | Leu | Ala 95 | Asp |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Lys | Leu | Pro | Pro 100 | Ile | Pro | Phe | Gly | Glu 105 | Ser | Val | Met | Asp | Leu 110 | Val | Val |
| Ile | Gly | Cys 115 | Gly | Pro | Ala | Gly | Leu 120 | Ser | Leu | Ala | Ala | Glu 125 | Ala | Ala | Lys |
| Leu | Gly 130 | Leu | Lys | Val | Gly | Leu 135 | Ile | Gly | Pro | Asp | Leu 140 | Pro | Phe | Thr | Asn |
| Asn 145 | Tyr | Gly | Val | Trp | Glu 150 | Asp | Glu | Phe | Lys | Asp 155 | Leu | Gly | Leu | Glu | Arg 160 |
| Cys | Ile | Glu | His | Ala 165 | Trp | Lys | Asp | Thr | Ile 170 | Val | Tyr | Leu | Asp | Asn 175 | Asp |
| Ala | Pro | Val | Leu 180 | Ile | Gly | Arg | Ala | Tyr 185 | Gly | Arg | Val | Ser | Arg 190 | His | Leu |
| Leu | His | Glu 195 | Glu | Leu | Leu | Lys | Arg 200 | Cys | Val | Glu | Ser | Gly 205 | Val | Ser | Tyr |
| Leu | Asp 210 | Ser | Lys | Val | Glu | Arg 215 | Ile | Thr | Glu | Ala | Gly 220 | Asp | Gly | His | Ser |
| Leu 225 | Val | Val | Cys | Glu | Asn 230 | Glu | Ile | Phe | Ile | Pro 235 | Cys | Arg | Leu | Ala | Thr 240 |
| Val | Ala | Ser | Gly | Ala 245 | Ala | Ser | Gly | Lys | Leu 250 | Leu | Glu | Tyr | Glu | Val 255 | Gly |
| Gly | Pro | Arg | Val 260 | Cys | Val | Gln | Thr | Ala 265 | Tyr | Gly | Val | Glu | Val 270 | Glu | Val |
| Glu | Asn | Asn 275 | Pro | Tyr | Asp | Pro | Asn 280 | Leu | Met | Val | Phe | Met 285 | Asp | Tyr | Arg |
| Asp | Tyr 290 | Met | Gln | Gln | Lys | Leu 295 | Gln | Cys | Ser | Glu | Glu 300 | Glu | Tyr | Pro | Thr |
| Phe 305 | Leu | Tyr | Val | Met | Pro 310 | Met | Ser | Pro | Thr | Arg 315 | Leu | Phe | Phe | Glu | Glu 320 |
| Thr | Cys | Leu | Ala | Ser 325 | Lys | Asp | Ala | Met | Pro 330 | Phe | Asp | Leu | Leu | Lys 335 | Arg |
| Lys | Leu | Met | Ser 340 | Arg | Leu | Lys | Thr | Leu 345 | Gly | Ile | Gln | Val | Thr 350 | Lys | Val |
| Tyr | Glu | Glu 355 | Glu | Trp | Ser | Tyr | 11e 360 | Pro | Val | Gly | Gly | Ser 365 | Leu | Pro | Asn |
| Thr | Glu 370 | Gln | Lys | Asn | Leu | Ala 375 | Phe | Gly | Ala | Ala | Ala 380 | Ser | Met | Val | His |
| Pro 385 | Ala | Thr | Gly | Tyr | Ser 390 | Val | Val | Arg | Ser | Leu 395 | Ser | Glu | Ala | Pro | Lys 400 |
| Tyr | Ala | Ser | Val | Ile | Ala | Lys | lle | Leu | Lys | Gln | Asp | Asn | Ser | Ala | Tyr |



Val Val Ser Gly Gln Ser Ser Ala Val Asn Ile Ser Met Gln Ala Trp 420 425

Ser Ser Leu Trp Pro Lys Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu 440

Phe Gly Leu Glu Leu Ile Val Gln Leu Asp Ile Glu Ala Thr Arg Thr 455

Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp Trp Gly Phe 470

Leu Gly Ser Ser Leu Ser Ser Phe Asp Leu Val Leu Phe Ser Met Tyr 490

Met Phe Val. Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His 500 505

Leu Leu Ser Asp Pro Ser Gly Ala Val Met Val Arg Ala Tyr Leu Glu 525

Arg

<210> 24 <211> 1370

<212> DNA

<213> Potato

<400> 24

BRIGHTON - WO GOSTOREAT I

TAGCGGAGGA TGAGTTCAAA GATCTTGGTC TTCAAGCCTG CATTGAACAT GTTTGGCTGG 60 GATACCATTG TATATCTTGA TGATGATGAT CCTATTCTTA TTGGCCGTGC CTATGGAAGA 120 GTTAGTCGCC ATTTACTGCA CGAGGAGTTA CTCAAAAGGT GTGTGGAGGC AGGTGTTTTG 180 TATCTAAACT CGAAAGTGGA TAGGATTGTT GAGGCCACAA ATGGCCACAG TCTTGTAGAG 240 TGCGAGGGTG ATGTTGTGAT TCCCTGCAGG TTTGTGACTG TTGCATCGGG AGCAGCCTCG 300 GGGAAATTCT TGCAGTATGA GTTGGGAGGT CCTAGAGTTT CTGTTCAAAC AGCTTATGGA 360 GTGGAAGTTG AGGTCGATAA CAATCCATTT GACCCGAGCC TGATGGTTTT CATGGATTAT 420 AGAGACTATG TCAGACACGA CGCTCAATCT TTAGAAGCTA AATATCCAAC ATTTCTCTAT 480 GCCATGCCCA TGTCTCCAAC ACGAGTCTTT TTCGAGGAAA CTTGTTTGGC TTCAAAAGAT 540 GCAATGCCAT TCGATCTGTT AAAGAAAAA TTGATGTTAC GATTGAACAC CCTCGGTGTA 600 AGAATTAAAG AAATTTATGA GGAGGAATGG TCTTACATAC CAGTTGGAGG ATCTTTGCCA 660 AATACAGAAC AAAAAACACT TGCATTTGGT GCTGCTGCTA GCATGGTTCA TCCAGCCACA 720 GGTTATTCAG TCGTCAGATC ACTGTCTGAA GCTCCAAAAT GCGCCTTCGT GCTTGCAAAT 780 ATATTACGAC AAAATCATAG CAAGAATATG CTTACTAGTT CAAGTACCCC GAGTATTTCA 840 ACTCAAGCTT GGAACACTCT TTGGCCACAA GAACGAAAAC GACAAAGATC GTTTTTCCTA 900

| ' | | |
|-------------|--|----------------|
| WO 99/63055 | | PCT/US99/12121 |

| TTTGGACTGG CTCT | GATATT GCAGO | TGGAT AT | rgaggga | TAAGGTCAT | T TTTCCGCGCG | | | | |
|---|--------------------|----------------|----------------|-------------------|-------------------------------|--|--|--|--|
| TTCTTCCGTG TGCC | AAAATG GATGT | GGCAG GG | ATTTCTTG | GTTCAAGTC | T TTCTTAGCAG | | | | |
| ACCTCATGTT ATTT | GCCTTC TACAT | GTTTA TT | ATTGCACC | AAATGACAT | G AGAAGAGGCT | | | | |
| TAATCAGACA TCTT | TTATCT GATCO | TACTG GTO | GCAACATT | GATAAGAAC | T TATCTTACAT | | | | |
| TTTAGAGTAA ATTCCTCCTA CAATAGTTGT TGAAAGAGGC CTCATTACTT CAGATTCATA | | | | | | | | | |
| ACAGAAATCG CGGT | CTCTCG AGGCC | TTGTA TAT | TAACATTT | TCACTAGGT | T AATATTGCTT | | | | |
| GAATAAGTTG CACA | GTTTCA GTTTT | TGTAT CTO | SCTTCTTT | TTTGTCCAA | G ATCATGTATT | | | | |
| GACCAATTTA TATA | CATTGC CAGTA | AAA TATAT. | ATATTTTA | AAAAAAAA | A | | | | |
| <210> 25 <211> 377 <212> PRT <213> Potato | | | | | | | | | |
| <400> 25 | Aco Ion Clu | Ton Cla | 73 - 0 | 7 3 03 0 | | | | | |
| Asp Glu Phe Lys 1 | 5 | ren Gin | 10 | lle Glu H | is Val Trp 15 | | | | |
| Arg Asp Thr Ile 20 | Val.Tyr Leu | Asp Asp 25 | Asp Asp | | eu Ile Gly 30 | | | | |
| Arg Ala Tyr Gly 35 | Arg Val Ser | Arg His | Leu Leu | His Glu G | lu Leu Leu | | | | |
| Lys Arg Cys Val 50 | Glu Ala Gly 55 | | Tyr Leu | Asn Ser Ly 60 | ys Val Äsp | | | | |
| Arg Ile Val Glu 65 | Ala Thr Asn 70 | Gly His | Ser Leu 75 | Val Glu C | ys Glu Gly 80 | | | | |
| Asp Val Val Ile | Pro Cys Arg 85 | Phe Val | Thr Val 90 | Ala Ser G | ly Ala Ala ['] 95 | | | | |
| Ser Gly Lys Phe 100 | Leu Gln Tyr | Glu Leu 105 | Gly Gly | | al Ser Val 10 | | | | |
| Gln Thr Ala Tyr 115 | Gly Val Glu | Val Glu 120 | Val Asp | Asn Asn Pi 125 | co Phe Asp | | | | |
| Pro Ser Leu Met 130 | Val Phe Met 135 | Asp Tyr | Arg Asp | Tyr Val Ai | rg His Asp | | | | |
| Ala Gln Ser Leu 145 | Glu Ala Lys 150 | Tyr Pro | Thr Phe | Leu Tyr Al | la Met Pro 160 | | | | |
| Met Ser Pro Thr | Arg Val Phe 165 | Phe Glu | Glu Thr 170 | Cys Leu Al | la Ser Lys 175 | | | | |
| Asp Ala Met Pro 180 | Phe Asp Leu | Leu Lys 185 | Lys Lys | | eu Arg Leu 90 | | | | |
| Asn Thr Leu Gly 195 | Val Arg Ile | Lys Glu 200 | Ile Tyr | Glu Glu Gl 205 | lu Trp Ser | | | | |

PNSUCIU- -WU



Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Thr Leu 210 215 220

Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser 225 230 235 240

Val Val Arg Ser Leu Ser Glu Ala Pro Lys Cys Ala Phe Val Leu Ala 245 250 255

Asn Ile Leu Arg Gln Asn His Ser Lys Asn Met Leu Thr Ser Ser Ser 260 265 270

Thr Pro Ser Ile Ser Thr Gln Ala Trp Asn Thr Leu Trp Pro Gln Glu 275 280 285

Arg Lys Arg Gln Arg Ser Phe Phe Leu Phe Gly Leu Ala Leu Ile Leu 290 295 300

Gln Leu Asp Ile Glu Gly Ile Arg Ser Phe Phe Arg Ala Phe Phe Arg 305 310 315 320

Val Pro Lys Met Met Trp Gly Phe Leu Gly Ser Ser Leu Ser Xaa Ala 325 330 335

Asp Leu Met Leu Phe Ala Phe Tyr Met Phe Ile Ile Ala Pro Asn Asp 340 345 350

Met Arg Arg Gly Leu Ile Arg His Leu Leu Ser Asp Pro Thr Gly Ala 355 360 365

Thr Leu Ile Arg Thr Tyr Leu Thr Phe 370 375

<210> 26

<211> 533

<212> PRT

<213> Chimeric lettuce/potato

<400> 26

Met Glu Cys Phe Gly Ala Arg Asn Met Thr Ala Thr Met Ala Val Phe 1 5 10 15

Thr Cys Pro Arg Phe Thr Asp Cys Asn Ile Arg His Lys Phe Ser Leu 20 25 30

Leu Lys Gly Arg Arg Phe Thr Asn Leu Ser Ala Ser Ser Ser Leu Arg 35 40 45

Gln Ile Lys Cys Ser Ala Lys Ser Asp Arg Cys Val Val Asp Lys Gln
50 60

Gly Ile Ser Val Ala Asp Glu Glu Asp Tyr Val Lys Ala Gly Gly Ser 65 70 75 80

Glu Leu Phe Phe Val Gln Met Gln Arg Thr Lys Ser Met Glu Ser Gln 85 90 . 95

Ser Lys Leu Ser Glu Lys Leu Ala Gln Ile Pro Ile Gly Asn Cys Ile 100 105 110

Leu Asp Leu Val Val Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala

BNSDOCID: <WO 9963055A1 1 >

| | | 115 | | | | | 120 | | | | | 125 | | | .∓`i |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Ala | Glu 130 | Ser | Ala | Lys | Leu | GÌy 135 | Leu | Asn | Val | Gly | Leu 140 | Ile | Gly | Pro | Asp |
| Leu 145 | Pro | Phe | Thr | Asn | Asn 150 | Tyr | Gly | Val | Trp | Gln 155 | Asp | Glu | Phe | Ile | Gly 160 |
| Leu | Gly | Leu | Glu | Gly 165 | Cys | Ile | Glu | His | Ser 170 | Trp | Lys | Asp | Thr | Leu 175 | Val |
| Tyr | Leu | Asp | Asp 180 | Ala | Asp | Pro | Ile | Arg 185 | Ile | Gly | Arg | Ala | Tyr 190 | Gly | Arg |
| Val | His | Arg 195 | Asp | Leu | Leu | His | Glu 200 | Glu | Leu | Leu | Arg | Arg 205 | Cys | Val | Glu |
| Ser | Gly 210 | Val | Ser | Tyr | Leu | Ser 215 | Ser | Lys | Val | Glu | Arg 220 | Ile | Thr | Glu | Ala |
| Pro 225 | Asn | Gly | Tyr | Ser | Leu 230 | Ile | Glu | Суѕ | Glu | Gly 235 | Asn | Ile | Thr | Ile | Pro 240 |
| Cys | Arg | Leu | Ala | Thr 245 | Val | Ala | Ser | Gly | Ala 250 | Ala | Ser | Gly | Lys | Phe 255 | Leu |
| Glu | Tyr | Glu | Leu 260 | Gly | Gly | Pro | Arg | Val 265 | Ser | Val | Gln | Thr | Ala 270 | Tyr | Gly |
| Val | Glu | Val 275 | Glu | Val | Asp | Asn | Asn 280 | Pro | Phe | Asp | Pro | Ser 285 | Leu | Met | Val |
| Phe | Met 290 | Asp | Tyr | Arg | Asp | Tyr 295 | Val | Arg | His | Asp | Ala 300 | Gln | Ser | Leu | Glu |
| Ala 305 | Lys | Tyr | Pro | Thr | Phe 310 | Leu | Tyr | Ala | Met | Pro 315 | Met | Ser | Pro | Thr | Arg 320 |
| Val | Phe | Phe | Glu | Glu 325 | Thr | Cys | Leu | Ala | Ser 330 | Lys | Asp | Ala | Met | Pro 335 | Phe |
| Asp | Leu | Leu | Lys 340 | Lys | Lys | Leu | Met | Leu 345 | Arg | Leu | Asn | Thr | Leu 350 | Gly | Val |
| Arg | Ile | Lys 355 | Glu | Ile | Tyr | Glu | Glu 360 | Glu | Trp | Ser | Tyr | Ile 365 | Pro | Val | Gly |
| Gly | Ser 370 | Leu | Pro | Asn | Thr | Glu 375 | Gln | Lys | Thr | Leu | Ala 380 | Phe | Gly | Ala | Ala |
| Ala 385 | Ser | Met | Val | His | Pro 390 | Ala | Thr | Gly | Tyr | Ser 395 | Val | Val | Arg | Ser | Leu 400 |
| Ser | Glu | Ala | Pro | Lys 405 | Cys | Ala | Phe | Val | Leu 410 | Ala | Asn | Ile | Leu | Arg 415 | Gln |
| Asn | His | Ser | Lys 420 | Asn | Met | Leu | Thr | Ser 425 | Ser | Ser | Thr | Pro | Ser 430 | Ile | Ser |
| Thr | Gln | Ala 435 | Trp | Asn | Thr | Leu | Trp 440 | Pro | Gln | Glu | Arg | Lys 445 | Arg | Gln | Arg |
| | | | | | | | | | | | | | | | |

WO 99/63055

Ser Phe Phe Leu Phe Gly Leu Ala Leu Ile Leu Gln Leu Asp Ile Glu 450 460

Gly Ile Arg Ser Phe Phe Arg Ala Phe Phe Arg Val Pro Lys Trp Met 465 470 475 480

Trp Gln Gly Phe Leu Gly Ser Ser Leu Ser Xaa Ala Asp Leu Met Leu 485 490 495

Phe Ala Phe Tyr Met Phe Ile Ile Ala Pro Asn Asp Met Arg Arg Gly 500 505

Leu Ile Arg His Leu Leu Ser Asp Pro Thr Gly Ala Thr Leu Ile Arg 515 520 525

Thr Tyr Leu Thr Phe 530

<210> 27

<211> 374

<212> PRT

<213> Arabidopsis thaliana

<400> 27

Glu Asp Glu Phe Asn Asp Leu Gly Leu Gln Lys Cys Ile Glu His Val
1 5 10 15

Trp Arg Glu Thr Ile Val Tyr Leu Asp Asp Asp Lys Pro Ile Thr Ile 20 25 30

Gly Arg Ala Tyr Gly Arg Val Ser Arg Arg Leu Leu His Glu Glu Leu 35 40 45

Leu Arg Arg Cys Val Glu Ser Gly Val Ser Tyr Leu Ser Ser Lys Val
50 60

Asp Ser Ile Thr Glu Ala Ser Asp Gly Leu Arg Leu Val Ala Cys Asp 65 70 75 80

Asp Asn Asn Val Ile Pro Cys Arg Leu Ala Thr Val Ala Ser Gly Ala 85 90 95

Ala Ser Gly Lys Leu Leu Gln Tyr Glu Val Gly Gly Pro Arg Val Cys 100 105 110

Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Glu Asn Ser Pro Tyr 115 120 125

Asp Pro Asp Gln Met Val Phe Met Asp Tyr Arg Asp Tyr Thr Asn Glu 130 135 140

Lys Val Arg Ser Leu Glu Ala Glu Tyr Pro Thr Phe Leu Tyr Ala Met 145 150 155 160

Pro Met Thr Lys Ser Arg Leu Phe Phe Glu Glu Thr Cys Leu Ala Ser 165 170 175

Lys Asp Val Met Pro Phe Asp Leu Leu Lys Thr Lys Leu Met Leu Arg 180 185 190

Leu Asp Thr Leu Gly Ile Arg Ile Leu Lys Thr Tyr Glu Glu Glu Trp

480

540

600

195 200 205 Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Asn 210 215 Leu Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr 235 Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys Tyr Ala Ser Val Ile 245 255 Ala Glu Ile Leu Arg Glu Glu Thr Thr Lys Gln Ile Asn Ser Asn Ile 260 Ser Arg Gln Ala Trp Asp Thr Leu Trp Pro Pro Glu Arg Lys Arg Gln 280 Arg Ala Phe Phe Leu Phe Gly Leu Ala Leu Ile Val Gln Phe Asp Thr Glu Gly Ile Arg Ser Phe Phe Arg Thr Phe Phe Arg Leu Pro Lys Trp 310 Met Trp Gln Gly Phe Leu Gly Ser Thr Leu Thr Ser Gly Asp Leu Val 325 Leu Phe Ala Leu Tyr Met Phe Val Ile Ser Pro Asn Asn Leu Arg Lys 345 Gly Leu Ile Asn His Leu Ile Ser Asp Pro Thr Gly Ala Thr Met Ile 355 Lys Thr Tyr Leu Lys Val 370 <210> 28 <211> 1002 <212> DNA <213> Adonis palaestina <400> 28 ATTCATCTTC AGCAGCGCTG TCGTACTCTT TCTATATCTT CTTCCATCAC TAACAGTAGT 60 CGCCGACGGT TGAATCGGCT ATTCGCCTCA ACGTCAACTA TGGGTGAAGT CACTGATGCT 120 GGAATGGATG CTGTTCAGAA GCGGCTCATG TTCGACGACG AATGTATTTT GGTGGATGAG 180 AATGACAAGG TCGTCGGGCA TGATTCCAAA TACAACTGTC ATTTGATGGA AAAGATAGAG 240 GCAGAAAATT TGCTTCACAG AGCCTTCAGT GTTTTCTTGT TCAACTCAAA ATATGAATTG 300 CTTCTTCAGC AACGATCCGC CACAAAGGTA ACATTCCCGC TCGTATGGAC AAACACATGT 360 TGCAGTCATC CTCTCTTTCG TGATTCCGAG CTCATAGAAG AAAATTATCT CGGTGTACGA

AACGCTGCAC AAAGAAAGCT TTTAGACGAG CTAGGCATTC CAGCTGAAGA TGTCCCAGTT

GATGAATTTA CTCCTCTTGG TCGCATTCTT TACAAAGCTC CATCTGACGG CAAATGGGGA

GAGCACGAAT TGGACTATCT CCTATTTATT GTCCGAGATG TGAAATACGA TCCAAACCCA

WO 99/63055 PCT/US99/12121

| • | | | | | | |
|---|--------------|------------|------------|------------|--------------|------|
| GATGAAGTT | G CTGATGCTAA | GTATGTTAA1 | CGCGAGGAG | TGAGAGAGA' | F ACTGAGAAAA | 660 |
| GCTGATGCT | GTGAAGAGGG | ACTCAAGTT(| TCTCCTTGGT | TTAGATTGG | T TGTTGATAAC | 720 |
| TTTTTGTTC | AGTGGTGGGA | TCATGTAGAG | CAGGGTACGA | TTAAGGAAG | T TGCTGACATG | 780 |
| AAAACTATCC | ACAAGTTGAC | TTAAGAGGAC | TTCTCTCCTC | TGTTCTACTA | TTTGTTTTT | 840 |
| GCTACAATAA | GTGGGTGGTG | ATAAGCAGTI | TTTCTGTTTT | CTTTAATTT | TGGCTTTTGA | 900 |
| ATTTGCCTCG | ATGTTGAACT | TGTAACATAT | TTAGACAAAT | ATGAGACCT1 | GTAAGTTGAA | 960 |
| TTTGAGGCTG | AATTTATATT | TTTGGGAACA | TAATAATGTT | AA | | 1002 |
| <210> 29 <211> 1271 <212> DNA <213> Adon | is palaesti | na | , | | | |
| <400> 29 | | | • | | | |
| TTTTAAAGCT | CTTTCGCTCC | ACCACCATCA | AAGCCAGCCA | AATTTCTCTG | TACAAAAGTT | 60 |
| | | | | | GATACGCATC | 120 |
| TAAACCAGTA | ATTCTCGGTT | TTAATTTGTT | TCCTAAATTA | GGCCCCTTTC | CGGAATCCCG | 180 |
| AGAATTATGT | CGTCGATCAG | GATTAATCCT | TTATATAGTA | TCTTCTCCAC | CACCACTAAA | 240 |
| | CTTCGTGTTC | | | | | 300 |
| TTCTATTTCT | TCTTCCATCA | CTAACAGTCC | TCGCCGAGGG | TTGAATCGGC | TGTTCGCCTC | 360 |
| AACGTCGACT | ATGGGTGAAG | TCGCTGATGC | TGGTATGGAT | GCCGTCCAGA | AGCGGCTTAT | 420 |
| GTTCGACGAT | GAATGTATTT | TGGTGGATGA | GAATGACAAG | GTCGTCGGAC | ATGATTCCAA | 480 |
| ATACAACTGT | CATTTGATGG | AAAAGATAGA | GGCAGAAAAC | TTGCTTCACA | GAGCCTTCAG | 540 |
| TGTTTTCTTA | TTCAACTCAA | AATACGAGTT | GCTTCTTCAG | CAACGATCTG | CAACGAAGGT | 600 |
| AACATTCCCG | CTCGTATGGA | CAAACACCTG | TTGCAGCCAT | CCCCTCTTCC | GTGATTCCGA | 660 |
| | GAAAATTTTC | | | | | 720 |
| | CCAGCTGAAG | | • | | | 780 |
| | CCATCTGACG | | | | | 840 |
| TGTCCGAGAT | GTGAAATACG | ATCCAAACCC | AGATGAAGTT | GCTGACGCTA | AGTACGTTAA | 900 |
| TCGCGAGGAG | TTGAAAGAGA | TACTGAGAAA | AGCTGATGCA | GGTGAAGAGG | GAATAAAGTT | 960 |
| GTCTCCTTGG | TTTAGATTGG | TTGTGGATAA | CTTTTTGTTC | AAGTGGTGGG | ATCATGTAGA | 1020 |
| GGAGGGGAAG | ATTAAGGACG | TCGCCGACAT | GAAAACTATC | CACAAGTTGA | CTTAAGAGAA | 1080 |
| AGTCTCTTAA | GTTCTACTAT | TTGGTTTTTG | CTTCAATAAG | TGGATGGTGA | TGAGCAGTTT | 1140 |
| TTATGCTTCC | TTTAATTTTG | GCTTTTCAAT | TTGCTTTATG | TGTTGAACTT | GTAACATATT | 1200 |
| TAGTCAAATA | TGAGACCTTG | TGAGTTGAAT | TTGAGGTTAT | ATTTATAGTT | TTGGGAACAT | 1260 |
| • | | | | | | |

÷ . AAAAAAAA A 1271 <210> 30 <211> 1109 <212> DNA <213> Haematococcus pluvialis <400> 30 TGGAACCTGG CCCGGCGGCA GTCCGATGCC GCGATGCTTC GTTCGTTGCT CAGAGGCCTC 60 ACGCATATCC CGCGCGTGAA CTCCGCCCAG CAGCCCAGCT GTGCACACGC GCGACTCCAG 120 TTTAAGCTCA GGAGCATGCA GCTGCTTGCC GAGGACCGCA CAGACCACAT GAGGGGTGCA 180 AGCACCTGGG CAGGCGGCA GTCGCAGGAT GAGCTGATGC TGAAGGACGA GTGCATCTTA 240 GTGGATGCTG ACGACAACAT CACAGGCCAT GCCAGCAAGC TGGAGTGCCA CAAATTCCTA 300 CCACATCAGC CTGCAGGCCT GCTGCACCGG GCCTTCTCTG TGTTCCTGTT TGACGACCAG 360 GGGCGACTGC TGCTGCAACA GCGTGCACGC TCAAAAATCA CCTTCCCAAG TGTGTGGACG 420 AACACCTGCT GCAGCCACCC TCTACATGGG CAGACCCCAG ATGAGGTGGA CCAACTAAGC 480 CAGGTGGCCG ACGGCACAGT ACCTGGCGCA AAAGCTGCTG CCATCCGCAA GTTGGAGCAC 540 GAGCTGGGGA TACCAGCGCA CCAGCTGCCG GCAAGCGCGT TTCGCTTCCT CACGCGTTTG 600 CACTACTGTG CCGCGGACGT GCAGCCGGCT GCGACACAAT CAGCGCTCTG GGGCGAGCAC 660 GAGATGGACT ACATCTTATT CATCCGGGCC AACGTCACCT TGGCGCCCAA CCCTGACGAG 720 GTGGACGAAG TCAGGTACGT GACGCAAGAG GAGCTGCGGC AGATGATGCA GCCGGACAAC 780 GGGTTGCAAT GGTCGCCGTG GTTTCGCATC ATCGCCGCGC GCTTCCTTGA GCGTTGGTGG 840 GCTGACCTGG ACGCGGCCCT AAACACTGAC AAACACGAGG ATTGGGGAAC GGTGCATCAC 900 ATCAACGAAG CGTGAAGGCA GAAGCTGCAG GATGTGAAGA CACGTCATGG GGTGGAATTG 960 CGTACTTGGC AGCTTCGTAT CTCCTTTTTC TGAGACTGAA CCTGCAGAGC TAGAGTCAAT 1020 GGTGCATCAT ATTCATCGTC TCTCTTTTGT TTTAGACTAA TCTGTAGCTA GAGTCACTGA 1080 TGAATCCTTT ACAACTTTCA AAAAAAAA 1109 <210> 31 <211> 985 <212> DNA <213> Lactuca sativa <400> 31 TGCCAAAATG TTGAAATTTC CCCCTTTTAA AACCATTGCT ACCATGATCT CTTCTCCATA 60 TTCTTCCTTC TTGCTGCCTC GGAAATCTTC TTTCCCTCCA ATGCCGTCTC TCGCAGCCGC 120 TAGTGTTTTC CTCCACCCTC TTTCGTCTGC CGCTATGGGC GATTCCAGCA TGGATGCTGT 180 CCAGCGACGT CTCATGTTCG ATGACGAATG CATTTTGGTG GATGAGAATG ACAAAGTGGT 240 TGGCCATGAT ACTAAATACA ATTGTCATTT GATGGAGAAG ATTGAAAAGG GAAATATGCT 300

PCT/US99/12121

WO 99/63055

| WO | | | |
|----|--|--|--|
| | | | |



PCT/US99/12121

| | 73/ | | | | | | |
|-----------------------|--------------|------------|------------|------------|------------|-----|------------------|
| ACACAGAGCA | A TTCAGTGTGT | TCTTGTTCAA | CTCGAAATAT | GAATTACTCC | TTCAGCAACG | ; | 360 [°] |
| TTCTGCAACO | AAGGTGACTT | TCCCTTTGGT | ATGGACAAAC | ACGTGTTGCA | GCCATCCACT | | 420 |
| ATACAGGGAG | AGTGAGCTTA | TTGACGAAAA | CGCCCTTGGG | GTGAGGAATG | CTGCACAGAG | 4 | 480 |
| GAAGCTCCTC | GATGAACTCG | GCATCCCTGG | AGCAGATGTT | CCGGTTGATG | AGTTCACTCC | į | 540 |
| ATTGGGTCGC | ATTCTATACA | AGGCCGCATC | GGATGGAAAG | TGGGGAGAAC | ATGAACTTGA | (| 600 |
| TTACCTGCT | TTTATGGTAC | GTGATGTTGG | TTTGGATCCG | AACCCAGATG | AAGTGAAAGA | • | 660 |
| TGTAAAATAT | GTGAACCGGG | AAGAGCTGAA | GGAATTGGTA | AGGAAGGCGG | ATGCTGGTGA | - | 720 |
| AGAGGGTGTG | AAGCTGTCCC | CGTGGTTCAA | ATTGATTGTC | GATAATTTCT | TGTTTCAGTG | 7 | 780 |
| GTGGGATCGA | CTCCATAAGG | GAACCCTAAC | CGAAGCTATT | GATATGAAAA | CAATCCACAA | | 340 |
| ACTCACATA | AAACACTACA | CTAGTAGGAG | AGAGGATTAT | ATGAGATATT | TGTTATATGT | 9 | 900 |
| GAAATTGAAA | TTCAGATGAA | TGCTTGTÄTT | TATTTCTATT | TGGACAAACT | TCAACTTCTT | 9 | 960 |
| TTTGCTACCT | TATCAGAAAA | AAAAA | • . | | | 9 | 985 |
| <210> 32 <211> 988 | | • | | • | | | |
| <212> DNA | uca sativa | | | | | | |
| <400> 32 | uca sativa | | • | | | | |
| | AAAATCTCTT | CCATTAACTG | CTCAAATCTC | CACCTTCGCC | GGTCTTAATC | | 60 |
| TCCGCCGGCG | CACTTTCACC | ACCATAACCG | CCGCCATGGG | TGACGATTCC | GGCATGGACG | . 1 | 120 |
| CTGTCCAGAG | ACGTCTCATG | TTTGATGATG | AATGCATTTT | GGTTGATGAA | AATGACAATG | . 3 | 180 |
| TTCTTGGGCA | TGATACCAAA | TACAATTGTC | ACTTGATGGA | GAAGATTGAG | AAAGATAATT | 2 | 240 |
| TGCTTCATAG | AGCATTCAGT | GTATTTTAT | TCAATTCAAA | ATACGAATTA | CTCCTTCAGC | 3 | 300 |
| AAAGGTCAGA | AACCAAGGTG | ACATTTCCTT | TGGTATGGAC | AAACACCTGT | TGCAGCCATC | 3 | 360 |
| CACTATACAC | AGAATCGGAG | TTAATTCCCG | AAAATGCCCT | TGGGGTCAGA | AATGCTGCAC | 4 | 120 |
| AGAGGAAGCT | TCTAGATGAA | CTCGGTATCC | CTGCTGAAGA | TGTTCCAGTT | GATGAGTTCA | 4 | 180 |
| CAACTTTAGG | TCGCATGTTG | TACAAGGCTC | CATCTGATGG | AAAATGGGGT | GAACATGAAG | 5 | 540 |
| TTGATTACCT | ACTCTTCCTC | GTGCGTGACG | TTGCCGTGAA | CCCAAACCCT | GATGAGGTGG | 6 | 500 |
| CGGACATTAC | ATACGTGAAC | CAAGAAGAGT | TAAAAGAGTT | ACTAAGGAAG | GCGGATGCGG | 6 | 560 |
| GTGAGGAGGG | TTTGAAATTG | TCCCCATGGT | TTAGGCTAGT | GGTGGACAAC | TTCTTGTTCA | 7 | 720 |
| AATGGTGGGA | TCATGTCCAA | AAGGGGACAC | TCAATGAAGC | AATTGACATG | AAAACCATTC | 7 | 780 |
| ATAAGTTGAT | ATGAAAAATG | GTTAATATTT | ATGGTGGTGG | TTTGGAGCTA | ATAATTTGTG | 8 | 340 |
| TGTTCAAGTC | TCGGTCCTTC | TTTTTTTAAC | GTTTTTTTT | TTTCTTTTAT | TGGGAGTGTT | 9 | 900 |
| TATTGTGTAC | TTGTAACGTA | GGCCCTTTGG | TTACGCTTTA | AGAGTTTAAT | AAAGAACCAC | 9 | 960 |



| ССТТААТТТА ДАДАЛАЛАЛ ДАДАЛАЛА | e ju | 988 |
|--|------------|------------------|
| <210> 33 <211> 1874 <212> DNA <213> Chlamydomonas reinhardtii | | |
| <400> 33 GGCACGAGCT CGAGTTTGTT TTACCATGAC ATCGGGAATT TGGAAGCTTG | AACTACCTCA | 60 |
| ATTACTCAAG TAACTCGCGG CAACACATTT CGCGCGCCAT CGCTGTTTTC | TCTGCTCCAG | 120 [.] |
| CTACCGAGCA GCATTGCTTT AGATCGCTTT GATGTCATAA ACTCCCACTT | ATATGAGATC | 180 |
| CAGTTTCATC GAGCCCAAGC CCAGAGCGCA ACCTGTCTTA AGCCGCGGCA | GGGCGTCCAT | 240 |
| GCGCCTCGCG CAAAGCCGTG CTCTCGTTGC GCGTGTCAGC TCCGCCCTGT | GGCCGGGAGC | 300 |
| AGGACTTTCA CAGGCTCAAA GCGTTGCGGT GCGAATGGCG AGTTCGTCAA | CCTGGGAAGG | 360 |
| CACGGGCCTG AGCCAGGATG ACTTCATGCA GCGGGACGAG TGCTTGGTGG | TGGACGAGCA | 420 |
| GGACCGGCTG CTAGGCACCG CCAACAAGTA CGACTGCCAC CGCTTCGAGG | CGGCCAAGGG | 480 |
| CCAGCCCTGC GGCCGCCTGC ACCGCGCCTT CTCCGTGTTC CTGTTCAGCC | CCGACGGCCG | 540 |
| ACTGCTGCTG CAGCAGCGC CAGCCAGCAA GGTGACGTTC CCGGGTGTGT | GGACCAACAC | 600 |
| CTGCTGCTCG CACCCGCTGG CGGGCCAGGC GCCGGACGAG GTGGACCTGC | CGGCGGCGGT | 660 |
| AGCCTCGGGC CAGGTGCCGG GCATCAAGGC GGCGGCGGTG CGCAAGCTGC | AGCACGAGCT | 720 |
| GGGGATACCG CCGGAGCAGG TTCCCGCCTC CTCCTTCTCC TTCCTCACGC | GTCTGCACTA | 780 |
| CTGCGCCGCC GACACCGCCA CGCACGGCCC GGCGGCGGAG TGGGGCGAGC | ACGAGGTGGA | 840 |
| CTACGTGCTG TTCGTGCGGC CGCAGCAGCC CGTCAGCCTG CAGCCCAACC | CAGACGAGGT | 900 |
| GGACGCCACG CGCTACGTGA CGCTGCCGGA GCTTCAGTCC ATGATGGCGG | ACCCCGGCCT | 960 |
| CAGCTGGAGC CCCTGGTTCC GCATCCTGGC CACACAGCCC GCCTTCCTGC | CCGCCTGGTG | 1020 |
| GGGCGACCTG AAGCGGCGCT GGCGCCCGGG CGGCAGCCGA CTGTCGGACT | GGGGCACCAT | 1080 |
| CCACCGCGTC ATGTGAAGAA AAAGGGGAAG CAGGGGCGGG AGCGGGGGAT | GAATGGGAAT | 1140 |
| GTGAATGCGA TTGTGATGCG GCGTGGGATG AGGTCTGAAG ACAGGGGGAA | AATCGGGGGG | 1200 |
| CGGGCGTGAG CGTGTGTGTA CGTGAGCGAC AAAGCCGGGA GGCGGACCGC | GCGATGGGTA | 1260 |
| CATGTGTGT CGGAGGGTCG GTGGGTCGGT CGGTTGCGCG GCATAGCGTG | TTGTGTGTGT | 1320 |
| GCGGCTGCAG GGGTATGTGG GCACCCGGGC ACGGAGGAGA AGGCACACGC | AGGTGGCGCG | 1380 |
| GAGGTGTGTC AGGGGCCATG GGCGGGCCTC ACTCCTGGTC GTGCCCAGTG | GTCTCGTGGG | 1440 |
| CAGAGTGGCA GGGGCTGCAC CCATATGAGC GGCGCACTGC CGCGCTGGGC | TAAGTCCTTA | 1500 |
| TCACTTGGTG AGGTGGGCG AGGTGGCTGT GGGCGGCGGG CGCAGTGGCA | GAAGGACACG | 1560 |
| GTGTGTGAGC GGTGGAGCTC TGGCCGTGCC GGCCGTGAGG GGCGGATAGC | GATATGACGT | 1620 |
| | | |

WO 99/63055 PCT/US99/12121

| | | | | | | .,00,,,,121 |
|--|------------|------------|------------|------------|------------|-------------|
| TGTGCTTGGC | CGCTGTAATG | CGGGAGAATG | TGCAGGCCGC | GAGAAGCGGG | CGGTGGCAGG | 1680 |
| AGGCCGCAGG | CTGCAGCACC | CGTTGGGGAG | GTGCCACCTG | CAGGCGCGGC | GCCGGGCGGG | 1740 |
| CCTGAGTAAT | GGGCGCCTGA | GTAGTGGCGG | CCACAGGAGG | CGCAGGAGGC | AGCAGCAGGA | 1800 |
| GGACGAGCTG | GAGGGACCCG | TTGGCAACCC | AAGGTTGCGC | GTGTAACATA | GTGGCCATAC | 1860 |
| ААААААААА | AAAA | | | | | 1874 |
| <210> 34 <211> 954 <212> DNA <213> Taget | es erecta | | | | | |
| <400> 34 CCAAAAACAA | CTCAAATCTC | CTCCGTCGCT | CTTACTCCGC | CATGGGTGAC | GACTCCGGCA | 60 |
| TGGATGCTGT | TCAGCGACGT | CTCATGTTTG | ACGATGAATG | CATTTTGGTG | GATGAGTGTG | 120 |
| ACAATGTGGT | GGGACATGAT | ACCAAATACA | ATTGTCACTT | GATGGAGAAG | ATTGAAACAG | 180 |
| GTAAAATGCT | GCACAGAGCA | TTCAGCGTTT | TTCTATTCAA | TTCAAAATAC | GAGTTACTTC | 240 |
| TTCAGCAACG | GTCTGCAACC | AAGGTGACAT | TTCCTTTAGT | ATGGACCAAC | ACCTGTTGCA | 300 |
| GCCATCCACT | CTACAGAGAA | TCCGAGCTTG | TTCCCGAAAA | CGCCCTTGGA | GTAAGAAATG | 360 |
| CTGCACAGAG | GAAGCTGTTG | GATGAACTCG | GTATCCCTGC | TGAAGATGTT | CCCGTTGATC | 420 |
| AGTTTACTCC | TTTAGGTCGC | ATGCTCTACA | AGGCTCCATC | TGATGGAAAG | TGGGGAGAAC | 480 |
| ATGAACTTGA | CTACCTACTT | TTCATAGTGA | GAGACGTTGC | TGTAAACCCG | AACCCAGATG | 540 |
| AAGTGGCGGA | TATCAAATAT | GTGACCAGAA | GAGTTAAAGG | AGCTGCTAAG | GAAAGCAGAT | 600 |
| GCGGGGGAGG | AGGGTTTGAA | GCTGTCTCCA | TGGTTCAGGT | TAGTGGTTGA | TAACTTCTTG | 660 |
| TTCAAGTGGT | GGGATCATGT | GCAAAAGGGT | ACACTCACTG | AAGCAATTGA | TATGAAAACC | 720 |
| ATACACAAGC | TGATATAGAA | ACACACCCTC | AACCGAAAAG | TTCAAGCCTA | ATAATTCGGG | 780 |
| TTGGGTCGGG | TCTACCATCA | ATTGTTTTTT | TCTTTTAAGA | AGTTTTAATC | TCTATTTGAG | 840 |
| CATGTTGATT | CTTGTCTTTT | GTGTGTAAGA | TTTTGGGTTT | CGTTTCAGTT | GTAATAATGA | 900 |
| ACCATTGATG | GTTTGCAATT | TCAAGTTCCT | ATCGACATGT | AGTGATCTAA | AAAA | 954 |
| <210> 35 <211> 1031 <212> DNA <213> Oryza | a sativa | | | | · | |
| <400> 35 | CCTCCCCTC | ACCCCCCCC | CCCMMCMCCC | 000007 | | |
| | | AGGCGGCCGC | | | | 60 |
| | | CGGGATGGAC | | | | 120 |
| | | ACAAGACAAT | | | • | 180 |
| CATCIGATGG | AAAAAATCGA | ATCTGAAAAT | CTACTTCATA | GGGCTTTCAG | TGTATTCCTG | 240 |



| | 00000 | | | ~ | | C1/05///12 |
|------------------------------------|------------|------------|------------|------------|------------|------------|
| TTCAACTCAA | AATATGAACT | CCTACTCCAG | CAACGATCTG | CAACAAAGGT | TACATTTCC | т 300 |
| CTAGTTTGGA | CCAACACTTG | CTGCAGCCAT | CCTCTGTACC | GTGAGTCTGA | GCTTATACA | G 360 |
| GAAAACTACC | TTGGTGTTAG | AAATGCTGCT | CAGAGGAAGC | TCTTGGATGA | GCTGGGCAT | C 420 |
| CCAGCTGAAG | ATGTGCCAGT | TGACCAATTC | ACCCCTCTTG | GTCGGATGCT | TTACAAGGC | C 480 |
| CCATCTGATG | GAAAATGGGG | TGAACACGAG | CTTGACTACC | TGCTGTTCAT | CGTCCGCGA | 540 |
| GTGAAGGTAG | TCCCGAACCC | GGACGAAGTG | GCCGATGTGA | AATACGTGAG | CCGTGAGCA | G 600 |
| CTGAAGGAGC | TCATCCGCAA | AGCGGACGCC | GGAGAGGAAG | GCCTGAAGCT | GTCTCCCTG | G 660 |
| TTCCGGCTGG | TTGTTGACAA | CTTCCTCATG | GGCTGGTGGG | ATCACGTCGA | GAAAGGCAC | 720 |
| CTCAACGAGG | CCGTGGACAT | GGAGACCATC | CACAAGCTGA | AGTAAGGACT | GCGATGTTG | r 780 |
| GGCTGGAAAG | AATGATCCTG | AAGACTCTGT | TCTTGTGCTG | CTGCATATTA | CTCTTACCA | 3 840 |
| GGAAGTTGCA | GAAGTCAGAA | GAAGCTTTTG | TATGTTTCTG | GGTTTGGAGC | TTGGAAGTG | r 900 |
| TGGGCTCTGC | TGACTGAGAG | ATTCCCTTAT | AGAGTGTCTA | TGTTAATTTA | GCAAACTTC' | r 960 |
| ATATTATACA | TGATTAGTTA | ATTGTTCGGT | GTCTGAATAA | AGAACAATAG | CATGTTCCA | r 1020 |
| GTTTATTTGC | T | | | | | 1031 |
| <210> 36 <211> 232 <212> PRT | voa avaata | | | | | |

<213> Tagetes erecta

<400> 36

Met Gly Asp Asp Ser Gly Met Asp Ala Val Gln Arg Arg Leu Met Phe

Asp Asp Glu Cys Ile Leu Val Asp Glu Cys Asp Asn Val Val Gly His

Asp Thr Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Thr Gly Lys

Met Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu

Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro Leu Val

Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser Glu Leu

Val Pro Glu Asn Ala Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu 100

Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp Gln Phe

Thr Pro Leu Gly Arg Met Leu Tyr Lys Ala Pro Ser Asp Gly Lys Trp 130 135



| Gly Glu His (| Asp 150 | Tyr | Leu | Leu | Ile 155 | Val | Arg | Asp | Val | Ala 160 |
|---------------|------------|-----|-----|-----|------------|-----|-----|-----|-----|------------|
| | | | | | | | | | | |

Val Asn Pro Asn Pro Asp Glu Val Ala Asp Ile Lys Tyr Val Ser His 165 170 175

Glu Glu Leu Lys Glu Leu Leu Arg Lys Ala Asp Ala Gly Glu Glu Gly 180 185 190

Leu Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe Leu Phe 195 200 205

Lys Trp Trp Asp His Val Gln Lys Gly Thr Leu Thr Glu Ala Ile Asp 210 215 220

Met Lys Thr Ile His Lys Leu Ile 225 230

<210> 37

<211> 280

<212> PRT

<213> Lactuca Sativa

<400> 37

Met Leu Lys Phe Pro Pro Phe Lys Thr Ile Ala Thr Met Ile Ser Ser 1 5 10 15

Pro Tyr Ser Ser Phe Leu Leu Pro Arg Lys Ser Ser Phe Pro Pro Met 20 25 30

Pro Ser Leu Ala Ala Ala Ser Val Phe Leu His Pro Leu Ser Ser Ala 35 40 45

Ala Met Gly Asp Ser Ser Met Asp Ala Val Gln Arg Arg Leu Met Phe 50 55 60

Asp Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Lys Val Val Gly His 65 70 75 80

Asp Thr Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Lys Gly Asn 85 90 95

Met Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu 100 105 110

Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro Leu Val 115 120 125

Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser Glu Leu 130 135 140

Ile Asp Glu Asn Ala Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu 145 150 155 160

Leu Asp Glu Leu Gly Ile Pro Gly Ala Asp Val Pro Val Asp Glu Phe 165 170 175

Thr Pro Leu Gly Arg Ile Leu Tyr Lys Ala Ala Ser Asp Gly Lys Trp
180 185 190

Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Met Val Arg Asp Val Gly

<u>.</u>.

195 . 200

Leu Asp Pro Asn Pro Asp Glu Val Lys Asp Val Lys Tyr Val Asn Arg 210 215 220

Glu Glu Leu Lys Glu Leu Val Arg Lys Ala Asp Ala Gly Glu Glu Gly 225 230 235 240

Val Lys Leu Ser Pro Trp Phe Lys Leu Ile Val Asp Asn Phe Leu Phe 245 250 255

Gln Trp Trp Asp Arg Leu His Lys Gly Thr Leu Thr Glu Ala Ile Asp 260 265 270

Met Lys Thr Ile His Lys Leu Thr 275 280

<210> 38

<211> 229

<212> PRT

<213> Lactuca Sativa

<400> 38

Met Gly Asp Asp Ser Gly Met Asp Ala Val Gln Arg Arg Leu Met Phe 1 5 10 15

Asp Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Asn Val Leu Gly His 20 25 30

Asp Thr Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Lys Asp Asn 35 40 45

Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu 50 55 60

Leu Leu Cln Gln Arg Ser Glu Thr Lys Val Thr Phe Pro Leu Val
65 70 75 80

Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser Glu Leu 85 90 95

Ile Pro Glu Asn Ala Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu 100 105 110

Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp Glu Phe 115 120 125

Thr Thr Leu Gly Arg Met Leu Tyr Lys Ala Pro Ser Asp Gly Lys Trp 130 135 140

Gly Glu His Glu Val Asp Tyr Leu Leu Phe Leu Val Arg Asp Val Ala 145 150 155 160

Val Asn Pro Asn Pro Asp Glu Val Ala Asp Ile Arg Tyr Val Asn Gln
165 170 175

Glu Glu Leu Lys Glu Leu Leu Arg Lys Ala Asp Ala Gly Glu Glu Gly 180 185

Leu Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe Leu Phe 195 200 205



Lys Trp Trp Asp His Val Gln Lys Gly Thr Leu Asn Glu Ala Ile Asp 210 215 220

Met Lys Thr Ile His 225

<210> 39

<211> 295

<212> PRT

<213> Adonis Palaestina

<400> 39

Met Ser Ser Ile Arg Ile Asn Pro Leu Tyr Ser Ile Phe Ser Thr Thr 1 5 10 15

Thr Lys Thr Leu Ser Ala Ser Cys Ser Ser Pro Ala Val His Leu Gln
20 25 30

Gln Arg Cys Arg Thr Leu Ser Ile Ser Ser Ser Ile Thr Asn Ser Pro 35 40 45

Arg Arg Gly Leu Asn Arg Leu Phe Ala Ser Thr Ser Thr Met Gly Glu
50 60

Val Ala Asp Ala Gly Met Asp Ala Val Gln Lys Arg Leu Met Phe Asp 65 70 75 80

Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Lys Val Val Gly Tyr Asp 85 90 95

Ser Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala Glu Asn Leu 100 105 110

Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu Leu 115 120 125

Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro Leu Val Trp 130 135 140

Thr Asn Thr Cys Cys Ser His Pro Leu Phe Arg Asp Ser Glu Leu Ile 145 150 155 160

Glu Glu Asn Phe Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu Leu 165 170 175

Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp Glu Phe Thr 180 185 190

Pro Leu Gly Arg Ile Leu Tyr Lys Ala Pro Ser Asp Gly Lys Trp Gly 195 200 205

Glu His Glu Leu Asp Tyr Leu Leu Phe Ile Val Arg Asp Val Lys Tyr 210 215 220

Asp Pro Asn Pro Asp Glu Val Ala Asp Ala Lys Tyr Val Asn Arg Glu 225 230 235 240

Glu Leu Lys Glu Ile Leu Arg Lys Ala Asp Ala Gly Glu Glu Gly Ile 245 250 . 255

Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe Leu Phe Lys

265

270

Trp Trp Asp His Val Glu Glu Gly Lys Ile Lys Asp Val Ala Asp Met 275 280 285

Lys Thr Ile His Lys Leu Thr 290 295

<210> 40

<211> 234

<212> PRT

<213> Adonis Palaestina

<400> 40

Met Gly Glu Val Thr Asp Ala Gly Met Asp Ala Val Gln Lys Arg Leu 1 5 10 15

Met Phe Asp Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Lys Val Val 20 25 30

Gly His Asp Ser Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala $35 \hspace{1cm} 40 \hspace{1cm} 45$

Glu Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys
50 55 60

Tyr Glu Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro 65 70 75 80

Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu Phe Arg Asp Ser 85 90 95

Glu Leu Ile Glu Glu Asn Tyr Leu Gly Val Arg Asn Ala Ala Gln Arg 100 105 110

Lys Leu Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp 115 120 125

Glu Phe Thr Pro Leu Gly Arg Ile Leu Tyr Lys Ala Pro Ser Asp Gly 130 135 140

Lys Trp Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Ile Val Arg Asp 145 150 155 160

Val Lys Tyr Asp Pro Asn Pro Asp Glu Val Ala Asp Ala Lys Tyr Val 165 170 175

Asn Arg Glu Glu Leu Arg Glu Ile Leu Arg Lys Ala Asp Ala Gly Glu 180 185 190

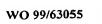
Glu Gly Leu Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe 195 200 205

Leu Phe Lys Trp Trp Asp His Val Glu Gln Gly Thr Ile Lys Glu Val 210 215 220

Ala Asp Met Lys Thr Ile His Lys Leu Thr 225 230

<210> 41 <211> 238

بيانيون



<212> PRT <213> Oryza Sativa

<400> 41

Met Ala Gly Ala Ala Ala Ala Val Glu Asp Ala Gly Met Asp Glu Val 1 5 10 15

Gln Lys Arg Leu Met Phe Asp Asp Glu Cys Ile Leu Val Asp Glu Gln
20 25 30

Asp Asn Val Val Gly His Glu Ser Lys Tyr Asn Cys His Leu Met Glu 35 40 45

Lys Ile Glu Ser Glu Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu 50 55 60

Phe Asn Ser Lys Tyr Glu Leu Leu Gln Gln Arg Ser Ala Thr Lys 65 70 75 / 80

Val Thr Phe Pro Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu 85 90 95

Tyr Arg Glu Ser Glu Leu Ile Gln Glu Asn Tyr Leu Gly Val Arg Asn 100 105 110

Ala Ala Gln Arg Lys Leu Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp 115 120 125

Val Pro Val Asp Gln Phe Thr Pro Leu Gly Arg Met Leu Tyr Lys Ala 130 135 140

Pro Ser Asp Gly Lys Trp Gly Glu His Glu Leu Asp Tyr Leu Leu Phe 145 150 155 160

Ile Val Arg Asp Val Lys Val Val Pro Asn Pro Asp Glu Val Ala Asp 165 170 175

Val Lys Tyr Val Ser Arg Glu Gln Leu Lys Glu Leu Ile Arg Lys Ala 180 185 190

Asp Ala Gly Glu Glu Leu Lys Leu Ser Pro Trp Phe Arg Leu Val 195 200 205

Val Asp Asn Phe Leu Met Gly Trp Trp Asp His Val Glu Lys Gly Thr 210 215 220

Leu Asn Glu Ala Val Asp Met Glu Thr Ile His Lys Leu Lys 225 230 235

<210> 42

<211> -233

<212> PRT

<213> Arabidopsis thaliana

<400> 42

Met Thr Asp Ser Asn Asp Ala Gly Met Asp Ala Val Gln Arg Arg Leu
1 5 10 15

Met Phe Glu Asp Glu Cys Ile Leu Val Asp Glu Asn Asn Arg Val Val 20 25 30

- Gly His Asp Thr Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Āla 35 40 45
- Glu Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys
 50 55 60
- Tyr Glu Leu Leu Gln Gln Arg Ser Lys Thr Lys Val Thr Phe Pro 65 70 75 80
- Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser 85 90 95
- Glu Leu Ile Glu Glu Asn Val Leu Gly Val Arg Asn Ala Ala Gln Arg 100 105 110
- Lys Leu Phe Asp Glu Leu Gly Ile Val Ala Glu Asp Val Pro Val Asp 115 120 125
- Glu Phe Thr Pro Leu Gly Arg Met Leu Tyr Lys Ala Pro Ser Asp Gly 130 135 140
- Lys Trp Gly Glu His Glu Val Asp Tyr Leu Leu Phe Ile Val Arg Asp 145 150 155 160
- Val Lys Leu Gln Pro Asn Pro Asp Glu Val Ala Glu Ile Lys Tyr Val 165 170 175
- Ser Arg Glu Glu Leu Lys Glu Leu Val Lys Lys Ala Asp Ala Gly Asp 180 185 190
- Glu Ala Val Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe 195 200 205
- Leu Met Lys Trp Trp Asp His Val Glu Lys Gly Thr Ile Thr Glu Ala 210 215 220
- Ala Asp Met Lys Thr Ile His Lys Leu 225 230
- <210> 43
- <211> 293
- <212> PRT
- <213> Haematococcus pluvialis
- <400> 43
- Met Leu Arg Ser Leu Leu Arg Gly Leu Thr His Ile Pro Arg Val Asn 1 5 10 15
- Ser Ala Gln Gln Pro Ser Cys Ala His Ala Arg Leu Gln Phe Lys Leu 20 25 30
- Arg Ser Met Gln Leu Leu Ser Glu Asp Arg Thr Asp His Met Arg Gly 35 40 45
- Ala Ser Thr Trp Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys
 50 55 60
- Asp Glu Cys Ile Leu Val Asp Val Glu Asp Asn Ile Thr Gly His Ala 65 70 75 80
- Ser Lys Leu Glu Cys His Lys Phe Leu Pro His Gln Pro Ala Gly Leu





Leu His Arg Ala Phe Ser Val Phe Leu Phe Asp Asp Gln Gly Arg Leu 100 105 110

Leu Leu Gln Gln Arg Ala Arg Ser Lys Ile Thr Phe Pro Ser Val Trp 115 120 125

Thr Asn Thr Cys Cys Ser His Pro Leu His Gly Gln Thr Pro Asp Glu 130 135 140

Val Asp Gln Leu Ser Gln Val Ala Asp Gly Thr Val Pro Gly Ala Lys
145 150 155 160

Ala Ala Ile Arg Lys Leu Glu His Glu Leu Gly Ile Pro Ala His 165 170 175

Gln Leu Pro Ala Ser Ala Phe Arg Phe Leu Thr Arg Leu His Tyr Cys 180 185 190

Ala Ala Asp Val Gln Pro Ala Ala Thr Gln Ser Ala Leu Trp Gly Glu 195 200 205

His Glu Met Asp Tyr Ile Leu Phe Ile Arg Ala Asn Val Thr Leu Ala 210 225 220

Pro Asn Pro Asp Glu Val Asp Glu Val Arg Tyr Val Thr Gln Glu Glu 225 235 240

Leu Arg Gln Met Met Gln Pro Asp Asn Gly Leu Gln Trp Ser Pro Trp 245 250 255

Phe Arg Ile Ile Ala Ala Arg Phe Leu Glu Arg Trp Trp Ala Asp Leu 260 265 270

Asp Ala Ala Leu Asn Thr Asp Lys His Glu Asp Trp Gly Thr Val His 275 280 285

His Ile Asn Glu Ala 290

<210> 44

<211> 304

<212> PRT

<213> Haematococcus pluvialis

<400> 44

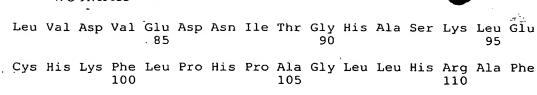
Met Leu Arg Ser Leu Leu Arg Gly Leu Thr His Ile Pro Arg Val Asn
1 5 10 15

Ser Ala Gln Gln Pro Ser Cys Ala His Ala Arg Leu Gln Phe Lys Leu 20 25 30

Arg Ser Met Gln Met Thr Leu Met Gln Pro Ser Ile Ser Ala Asn Leu 35 40 45

Ser Arg Ala Glu Asp Arg Thr Asp His Met Arg Gly Ala Ser Thr Trp 50 55 60

Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys Asp Glu Cys Ile 65 70 . 75 80



Ser Val Phe Leu Phe Asp Asp Gln Gly Arg Leu Leu Gln Gln Arg 115 120 125

Ala Arg Ser Lys Ile Thr Phe Pro Ser Val Trp Thr Asn Thr Cys Cys 130 135 140

Ser His Pro Leu His Gly Gln Thr Pro Asp Glu Val Asp Gln Leu Ser 145 150 155 160

Gln Val Ala Asp Gly Thr Val Pro Gly Ala Lys Ala Ala Ala Ile Arg 165 170 175

Lys Leu Glu His Glu Leu Gly Ile Pro Ala His Gln Leu Pro Ala Ser 180 185 190

Ala Phe Arg Phe Leu Thr Arg Leu His Tyr Cys Ala Ala Asp Val Gln
195 200 205

Pro Ala Ala Thr Gln Ser Ala Leu Trp Gly Glu His Glu Met Asp Tyr 210 220

Ile Leu Phe Ile Arg Ala Asn Val Thr Leu Ala Pro Asn Pro Asp Glu 225 230 235 240

Val Asp Glu Val Arg Tyr Val Thr Gln Glu Glu Leu Arg Gln Met Met 245 250 255

Gln Pro Asp Asn Gly Leu Gln Trp Ser Pro Trp Phe Arg Ile Ile Ala 260 265 270

Ala Arg Phe Leu Glu Arg Trp Trp Ala Asp Leu Asp Ala Ala Leu Asn 275 280 285

Thr Asp Lys His Glu Asp Trp Gly Thr Val His His Ile Asn Glu Ala 290 295 300

<210> 45

<211> 307

<212> PRT

<213> Chlamydomonas reinhardtii

<400> 45

Met Arg Ser Ser Phe Ile Glu Pro Lys Pro Arg Ala Gln Pro Val Leu 1 5 10 15

Ser Arg Gly Arg Ala Ser Met Arg Leu Ala Gln Ser Arg Ala Leu Val 20 25 30

Ala Arg Val Ser Ser Ala Leu Trp Pro Gly Ala Gly Leu Ser Gln Ala 35 40 45

Gln Ser Val Ala Val Arg Met Ala Ser Ser Ser Thr Trp Glu Gly Thr 50 55 60

Gly Leu Ser Gln Asp Asp Phe Met Gln Arg Asp Glu Cys Leu Val Val



65 70 .75 Asp Glu Gln Asp Arg Leu Leu Gly Thr Ala Asn Lys Tyr Asp Cys His Arg Phe Glu Ala Ala Lys Gly Gln Pro Cys Gly Arg Leu His Arg Ala Phe Ser Val Phe Leu Phe Ser Pro Asp Gly Arg Leu Leu Gln Gln 120 Arg Ala Ala Ser Lys Val Thr Phe Pro Gly Val Trp Thr Asn Thr Cys 130 135 Cys Ser His Pro Leu Ala Gly Gln Ala Pro Asp Glu Val Asp Leu Pro Ala Ala Val Ala Ser Gly Gln Val Pro Gly Ile Lys Ala Ala Ala Val 170 Arg Lys Leu Gln His Glu Leu Gly Ile Pro Pro Glu Gln Val Pro Ala Ser Ser Phe Ser Phe Leu Thr Arg Leu His Tyr Cys Ala Ala Asp Thr 200 Ala Thr His Gly Pro Ala Ala Glu Trp Gly Glu His Glu Val Asp Tyr 210 Val Leu Phe Val Arg Pro Gln Gln Pro Val Ser Leu Gln Pro Asn Pro 235 Asp Glu Val Asp Ala Thr Arg Tyr Val Thr Leu Pro Glu Leu Gln Ser Met Met Ala Asp Pro Gly Leu Ser Trp Ser Pro Trp Phe Arg Ile Leu 265 Ala Thr Gln Pro Ala Phe Leu Pro Ala Trp Trp Gly Asp Leu Lys Arg Arg Trp Arg Pro Gly Gly Ser Arg Leu Ser Asp Trp Gly Thr Ile His 295 Arg Val Met 305 <210> 46 <211> 1848 <212> DNA

<213> Adonis palaestina

<400> 46

| WO 99/ | 63055 | | | | F | CT/US99/12121 |
|------------|------------|------------|------------|------------|------------|---------------|
| TTTTCAAGTG | AGGGCTGATG | GTGGAAGCGG | GAGTAGAACT | TCTGTTGCTT | ATAAAGĀGG | G 300 |
| TTTTGTGGAC | GAGGAGGATT | TTATCAAAGC | TGGTGGTTCT | GAGCTTTTGT | TTGTCCAAA | т 360 |
| GCAGCAAACA | AAGTCTATGG | AGAAACAGGC | CAAGCTCGCC | GATAAGTTGC | CACCAATAC | C 420 |
| TTTCGGAGAA | TCTGTGATGG | ACTTGGTTGT | AATAGGTTGT | GGACCTGCTG | GTCTTTCAC | T 480 |
| GGCTGCAGAA | GCTGCTAAGC | TAGGCTTGAA | AGTTGGCCTT | ATTGGTCCTG | ATCTTCCTT | T 540 |
| TACAAATAAT | TATGGTGTGT | GGGAAGACGA | GTTCAAAGAT | CTTGGACTTG | AACGTTGTA | T 600 |
| CGAGCATGCT | TGGAAGGACA | CCATCGTATA | TCTTGACAAT | GATGCTCCTG | TCCTTATTG | G 660 |
| TCGTGCATAT | GGACGAGTTA | GCCGGCATTT | GCTGCATGAA | GAGTTGCTGA | AAAGGTGTG | T 720 |
| CGAGTCAGGT | GTATCATATC | TGAATTCTAA | AGTGGAAAGG | ATCACTGAAG | CTGGTGATG | G 780 |
| CCATAGTCTT | GTAGTTTGTG | AAAACGACAT | CTTTATCCCT | TGCAGGCTTG | CTACTGTTG | C 840 |
| ATCTGGAGCA | GCTTCAGGGA | AACTTTTGGA | GTATGAAGTA | GGTGGCCCTC | GTGTTTGTG | r 900 |
| CCAAACTGCT | TATGGTGTGG | AGGTTGAGGT | GGAGAACAAT | CCATACGATC | CCAACTTAA | r 960 |
| GGTATTTATG | GACTACAGAG | ACTATATGCA | ACAGAAATTA | CAGTGCTCGG | AAGAAGAAT. | A 1020 |
| TCCAACATTT | CTCTATGTCA | TGCCCATGTC | GCCAACAAGA | CTTTTTTTTG | AGGAAACCT | G 1080 |
| TTTGGCCTCA | AAAGATGCCA | TGCCTTTCGA | TCTACTGAAG | AGAAAACTAA | TGTCACGAT | r 1140 |
| GAAGACTCTG | GGTATCCAAG | TTACAAAAAT | TTATGAAGAG | GAATGGTCTT | ATATTCCTG | r 1200 |
| TGGGGGTTCT | TTACCAAACA | CAGAGCAAAA | GAACCTAGCA | TTTGGTGCTG | CAGCAAGCA' | г 1260 |
| GGTGCATCCA | GCAACAGGCT | ATTCGGTTGT | ACGATCACTA | TCAGAAGCTC | CAAAATATG | 2 1320 |
| TTCTGTAATT | GCAAAGATTT | TGAAGCAAGA | TAACTCTGCA | TATGTGGTTT | CTGGACAAA | 3 1380 |
| CAGTGCAGTA | AACATTTCAA | TGCAAGCATG | GAGCAGTCTT | TGGCCAAAGG | AGCGAAAAC | G 1440 |
| TCAAAGAGCA | TTCTTTCTTT | TCGGGTTAGA | GCTTATTGTG | CAGCTAGATA | TTGAAGCAA | 1500 |
| CAGAACGTTC | TTTAGAACCT | TCTTCCGCTT | GCCAACTTGG | ATGTGGTGGG | GTTTCCTTG | G 1560 |
| GTCTTCACTA | TCATCTTTCG | ATCTTGTATT | GTTTTCCATG | TACATGTTTG | TTTTGGCCC | 1620 |
| GAACAGCATG | AGGATGTCAC | TTGŢGAGACA | TTTGCTTTCA | GATCCTTCTG | GTGCAGTTA | r 1680 |
| GGTTAAAGCT | TACCTCGAAA | GGTAATCTGT | TTTATGAAAC | TATAGTGTCT | CATTAAATA | A 1740 |

Met Glu Leu Leu Gly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp

ATGAGGATCC TTCGTATATG TATATGATCA TCTCTATGTA TATCCTATAT TCTAATCTCA

TAAAGTAATC GAAAATTCAT TGATAGAAAA AAAAAAAAA AAAAAAAA

1800

1848

<210> 47

<211> 529

<212> PRT

<213> Adonis palaestina

<400> 47

15 Thr Phe Gly Thr Arg Asn Leu Ser Ser Ser Lys Leu Ala Tyr Asn Ile His Arg Tyr Gly Ser Ser Cys Arg Val Asp Phe Gln Val Arg Ala Asp Gly Gly Ser Gly Ser Arg Ser Ser Val Ala Tyr Lys Glu Gly Phe Val Asp Glu Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu Leu Leu Phe Val Gln Met Gln Gln Thr Lys Ser Met Glu Lys Gln Ala Lys Leu Ala Asp Lys Leu Pro Pro Ile Pro Phe Gly Glu Ser Val Met Asp Leu Val Val 105 Ile Gly Cys Gly Pro Ala Gly Leu Ser Leu Ala Ala Glu Ala Ala Lys Leu Gly Leu Lys Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Glu Arg 155 Cys Ile Glu His Ala Trp Lys Asp Thr Ile Val Tyr Leu Asp Asn Asp Ala Pro Val Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu His Glu Glu Leu Leu Lys Arg Cys Val Glu Ser Gly Val Ser Tyr Leu Asp Ser Lys Val Glu Arg Ile Thr Glu Ala Gly Asp Gly His Ser Leu Val Val Cys Glu Asn Glu Ile Phe Ile Pro Cys Arg Leu Ala Thr 225 Val Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Glu Tyr Glu Val Gly 250 Gly Pro Arg Val Cys Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Glu Asn Asn Pro Tyr Asp Pro Asn Leu Met Val Phe Met Asp Tyr Arg 280 Asp Tyr Met Gln Gin Lys Leu Gln Cys Ser Glu Glu Glu Tyr Pro Thr 295 Phe Leu Tyr Val Met Pro Met Ser Pro Thr Arg Leu Phe Phe Glu Glu 315 Thr Cys Leu Ala Ser Lys Asp Ala Met Pro Phe Asp Leu Leu Lys Arg 330



Lys Leu Met Ser Arg Leu Lys Thr Leu Gly Ile Gln Val Thr Lys Val 345 Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys 390 Tyr Ala Ser Val Ile Ala Lys Ile Leu Lys Gln Asp Asn Ser Ala Tyr Val Val Ser Gly Gln Ser Ser Ala Val Asn Ile Ser Met Gln Ala Trp

Ser Ser Leu Trp Pro Lys Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu

Phe Gly Leu Glu Leu Ile Val Gln Leu Asp Ile Glu Ala Thr Arg Thr

Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp Trp Gly Phe

Leu Gly Ser Ser Leu Ser Ser Phe Asp Leu Val Leu Phe Ser Met Tyr

Met Phe Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His 500 505

Leu Leu Ser Asp Pro Ser Gly Ala Val Met Val Arg Ala Tyr Leu Glu

Arg

<210> 48

<211> 378

<212> PRT

<213> Potato

<400> 48

Asp Glu Phe Lys Asp Leu Gly Leu Gln Ala Cys Ile Glu His Val Trp

Arg Asp Thr Ile Val Tyr Leu Asp Asp Asp Pro Ile Leu Ile Gly

Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu His Glu Glu Leu Leu

Lys Arg Cys Val Glu Ala Gly Val Leu Tyr Leu Asn Ser Lys Val Asp

Arg Ile Val Glu Ala Thr Asn Gly His Ser Leu Val Glu Cys Glu Gly

Asp Val Val Ile Pro Cys Arg Phe Val Thr Val Ala Ser Gly Ala Ala



| Ser Gly Lys Phe I 100 | Leu Gln Ty | yr Glu Leu 105 | Gly Gly Pro | Arg Val Ser Val |
|--------------------------|------------|-------------------|-------------|-----------------|
|--------------------------|------------|-------------------|-------------|-----------------|

Gln Thr Ala Tyr Gly Val Glu Val Glu Val Asp Asn Asn Pro Phe Asp 115 120 125

Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp Tyr Val Arg His Asp 130 135

Ala Gln Ser Leu Glu Ala Lys Tyr Pro Thr Phe Leu Tyr Ala Met Pro 145 150 155 160

Met Ser Pro Thr Arg Val Phe Phe Glu Glu Thr Cys Leu Ala Ser Lys 165 170 175

Asp Ala Met Pro Phe Asp Leu Leu Lys Lys Leu Met Leu Arg Leu 180 185 190

Asn Thr Leu Gly Val Arg Ile Lys Glu Ile Tyr Glu Glu Glu Trp Ser 195 200 205

Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Thr Leu 210 215 220

Ala Phe Gly Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser 225 230 235 240

Val Val Arg Ser Leu Ser Glu Ala Pro Lys Cys Ala Phe Val Leu Ala 245 250 255

Asn Ile Leu Arg Gln Asn His Ser Lys Asn Met Leu Thr Ser Ser Ser 260 265 270

Thr Pro Ser Ile Ser Thr Gln Ala Trp Asn Thr Leu Trp Pro Gln Glu 275 280 285

Arg Lys Arg Gln Arg Ser Phe Phe Leu Phe Gly Leu Ala Leu Ile Leu 290 295 300

Gln Leu Asp Ile Glu Gly Ile Arg Ser Phe Phe Arg Ala Phe Phe Arg 305 . 310 315 320

Val Pro Lys Trp Met Trp Gln Gly Phe Leu Gly Ser Ser Leu Ser Xaa 325 330 335

Ala Asp Leu Met Leu Phe Ala Phe Tyr Met Phe Ile Ile Ala Pro Asn 340 345 350

Asp Met Arg Arg Gly Leu Ile Arg His Leu Leu Ser Asp Pro Thr Gly 355 360 365

Ala Thr Leu Ile Arg Thr Tyr Leu Thr Phe 370

<210> 49

<211> 524

<212> PRT

<213> Arabidopsis thaliana

<400> 49

| Met 1 | Glu | Cys | Val | Gly 5 | Ala | Arg | Asn | Phe | Ala 10 | Ala | Met | Ala | Val | Ser 15 | Thr |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Phe | Pro | Ser | Trp 20 | Ser | Cys | Arg | Arg | Lys 25 | Phe | Pro | Val | Val | Lys 30 | Arg | Tyr |
| Ser | Tyr | Arg 35 | Asn | Ile | Arg | Phe | Gly 40 | Leu | Cys | Ser | Val | Arg 45 | Ala | Ser | Gly |
| Gly | Gly 50 | Ser | Ser | Gly | Ser | Glu 55 | Ser | Cys | Val | Ala | Val 60 | Arg | Glu | Asp | Phe |
| Ala 65 | Asp | Glu | Glu | Asp | Phe 70 | Val | Lys | Ala | Gly | Gly 75 | Ser | Glu | Ile | Leu | Phe 80 |
| Val | Gln | Met | Gln | Gln 85 | Asn | Lys | Asp | Met | Asp 90 | Glu | Gln | Ser | Lys | Leu 95 | Val |
| Asp | Lys | Leu | Pro 100 | Pro | Ile | Ser | Ile | Gly 105 | Asp | Gly | Ala | Leu | Asp 110 | His | Val |
| Val | Ile | Gly 115 | Cys | Gly | Pro | Ala | Gly 120 | Leu | Ala | Leu | Ala | Ala 125 | Glu | Ser | Ala |
| Lys | Leu 130 | Gly | Leu | Lys | Val | Gly 135 | Leu | Ile | Gly | Pro | Asp 140 | Leu | Pro | Phe | Thr |
| Asn 145 | Asn | Tyr | Gly | Val | Trp 150 | Glu | Asp | Glu | Phe | Asn 155 | Asp | Leu | Gly | Leu | Gln 160 |
| Lys | Cys | Ile | Glu | His 165 | Val | Trp | Arg | Glu | Thr 170 | Ile | Val | Tyr | Leu | Asp 175 | |
| Asp | Lys | Pro | Ile 180 | Thr | Ile | Gly | Arg | Ala 185 | Tyr | Gly | Arg | Val | Ser 190 | Arg | Arg |
| Leu | Leu | His 195 | Glu | Glu | Leu | Leu | Arg 200 | Arg | Cys | Val | Glu | Ser 205 | Gly | Val | Ser |
| Tyr | Leu 210 | Ser | Ser | Lys | Val | Asp 215 | Ser | Ile | Thr | Glu | Ala 220 | Ser | Asp | Gly | Leu |
| Arg 225 | Leu | Val | Ala | Cys | Asp 230 | Asp | Asn | Asn | Val | Ile 235 | Pro | Cys | Arg | Leu | Ala 240 |
| Thr | Val | Ala | Ser | Gly 245 | Ala | Ala | Ser | Gly | Lys 250 | Leu | Leu | Gln | Tyr | Glu 255 | Val |
| Gly | Gly | Pro | Arg 260 | Val | Cys | Val | Gln | Thr 265 | Ala | Tyr | Gly | Val | Glu 270 | Val | Glu |
| Val | Glu | Asn 275 | Ser | Pro | Tyr | Asp | Pro 280 | Asp | Gln | Met | Val | Phe 285 | Met | Asp | Tyr |
| Arg | Asp 290 | Tyr | Thr | Asn | Glu | Lys 295 | Val | Arg | Ser | Leu | Glu 300 | Ala | Glu | Tyr | Pro |
| Thr 305 | Phe | Leu | Tyr | Ala | Met 310 | Pro | Met | Thr | Lys | Ser 315 | Arg | Leu | Phe | Phe | Glu 320 |
| Glu | Thr | Cys | Leu | Ala | Ser | Lys | Asp | Val | Met | Pro | Phe | Asp | Leu | Leu | Lys |



335

325 330

Thr Lys Leu Met Leu Arg Leu Asp Thr Leu Gly Ile Arg Ile Leu Lys 340 345 350

Thr Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro 355 360 365

Asn Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val 370 380

His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro 385 390 395 400

Lys Tyr Ala Ser Val Ile Ala Glu Ile Leu Arg Glu Glu Thr Thr Lys
405 410 415

Gln Ile Asn Ser Asn Ile Ser Arg Gln Ala Trp Asp Thr Leu Trp Pro 420 425 430

Pro Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu Phe Gly Leu Ala Leu 435 440 445

Ile Val Gln Phe Asp Thr Glu Gly Ile Arg Ser Phe Phe Arg Thr Phe 450 455 460

Phe Arg Leu Pro Lys Trp Met Trp Gln Gly Phe Leu Gly Ser Thr Leu 465 470 475 480

Thr Ser Gly Asp Leu Val Leu Phe Ala Leu Tyr Met Phe Val Ile Ser 485 490 495

Pro Asn Asn Leu Arg Lys Gly Leu Ile Asn His Leu Ile Ser Asp Pro 500 505 510

Thr Gly Ala Thr Met Ile Lys Thr Tyr Leu Lys Val 515 520

<210> 50

<211> 529

<212> PRT

<213> Adonis palaestina

<400> 50

Met Glu Leu Leu Gly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp
1 5 10 15

Thr Phe Gly Thr Arg Asn Leu Ser Ser Ser Lys Leu Ala Tyr Asn Ile 20 25 30

His Arg Tyr Gly Ser Ser Cys Arg Val Asp Phe Gln Val Arg Ala Asp 35 40 45

Gly Gly Ser Gly Ser Arg Ser Ser Val Ala Tyr Lys Glu Gly Phe Val 50 60

Asp Glu Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu Leu Leu Phe Val 65 70 75 80

Gln Met Gln Gln Thr Lys Ser Met Glu Lys Gln Ala Lys Leu Ala Asp

PCT/US99/12121

90

95 Lys Leu Pro Pro Ile Pro Phe Gly Glu Ser Val Met Asp Leu Val Val 105 Ile Gly Cys Gly Pro Ala Gly Leu Ser Leu Ala Ala Glu Ala Ala Lys Leu Gly Leu Lys Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Glu Arg. Cys Ile Glu His Ala Trp Lys Asp Thr Ile Val Tyr Leu Asp Asn Asp Ala Pro Val Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu His Glu Glu Leu Leu Lys Arg Cys Val Glu Ser Gly Val Ser Tyr Leu Asp Ser Lys Val Glu Arg Ile Thr Glu Ala Gly Asp Gly His Ser Leu Val Val Cys Glu Asn Glu Ile Phe Ile Pro Cys Arg Leu Ala Thr 235 Val Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Glu Tyr Glu Val Gly Gly Pro Arg Val Cys Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val 260 Glu Asn Asn Pro Tyr Asp Pro Asn Leu Met Val Phe Met Asp Tyr Arg 280 Asp Tyr Met Gln Gln Lys Leu Gln Cys Ser Glu Glu Glu Tyr Pro Thr Phe Leu Tyr Val Met Pro Met Ser Pro Thr Arg Leu Phe Phe Glu Glu 315 Thr Cys Leu Ala Ser Lys Asp Ala Met Pro Phe Asp Leu Leu Lys Arg 330 Lys Leu Met Ser Arg Leu Lys Thr Leu Gly Ile Gln Val Thr Lys Val Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val His 375 Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys 395 Tyr Ala Ser Val Ile Ala Lys Ile Leu Lys Gln Asp Asn Ser Ala Tyr 410



Val Val Ser Gly Gln Ser Ser Ala Val Asn Ile Ser Met Gln Ala Trp
420 425 430

Ser Ser Leu Trp Pro Lys Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu 435 440 445

Phe Gly Leu Glu Leu Ile Val Gln Leu Asp Ile Glu Ala Thr Arg Thr 450 460

Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp Trp Gly Phe 465 470 475 480

Leu Gly Ser Ser Leu Ser Ser Phe Asp Leu Val Leu Phe Ser Met Tyr
485 490 495

Met Phe Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His 500 505 510

Leu Leu Ser Asp Pro Ser Gly Ala Val Met Val Arg Ala Tyr Leu Glu 515 520 525

Arg

<210> 51

<211> 529

<212> PRT

<213> Adonis palaestina

<400> 51

Met Glu Leu Leu Gly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp
1 5 10 15

Thr Phe Gly Thr Arg Asn Leu Ser Ser Ser Lys Leu Ala Tyr Asn Ile 20 25 30

His Arg Tyr Gly Ser Ser Cys Arg Val Asp Phe Gln Val Arg Ala Asp 35 40 45

Gly Gly Ser Gly Ser Arg Thr Ser Val Ala Tyr Lys Glu Gly Phe Val
50 60

Asp Glu Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu Leu Leu Phe Val 65 70 75 80

Gln Met Gln Gln Thr Lys Ser Met Glu Lys Gln Ala Lys Leu Ala Asp 85 90 95

Lys Leu Pro Pro Ile Pro Phe Gly Glu Ser Val Met Asp Leu Val Val 100 105 110

Ile Gly Cys Gly Pro Ala Gly Leu Ser Leu Ala Ala Glu Ala Ala Lys 115 120 125

Leu Gly Leu Lys Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn 130 135 140

Asn Tyr Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Glu Arg 145 150 155 160



| | | - | | | | | | | | | | | | | |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Cys | Ile | Glu | His | Ala 165 | Trp | Lys , | Asp | Thr | Ile 170 | Val | Tyr | Leu | Asp | Asn 175 | Asp |
| Ala | Pro | Val | Leu 180 | Ile | Gly | Arg | Ala | Tyr 185 | Gly | Arg | Val | Ser | Arg 190 | His | Leu |
| Leu | His | Glu 195 | Glu | Leu | Leu | Lys | Arg 200 | Cys | Val | Glu | Ser | Gly 205 | Val | Ser | Tyr |
| Leu | Asn 210 | Ser | Lys | Val | Glu | Arg 215 | Ile | Thr | Glu | Ala | Gly 220 | Asp | Gly | His | Ser |
| Leu 225 | Val | Val | Cys | Glu | Asn 230 | Asp | Ile | Phe | Ile | Pro 235 | Cys | Arg | Leu | Ala | Thr 240 |
| Val | Ala | Ser | Gly | Ala 245 | Ala | Ser | Gly | Lys | Leu 250 | Leu | Glu | Tyr | Glu | Val 255 | Gly |
| Gly | Pro | Arg | Val 260 | Cys | Val | Gln | Thr | Ala 265 | Tyr | Gly | Val | Glu | Val 270 | Glu | Val |
| Glu | Asn | Asn 275 | Pro | Tyr | Asp | Pro | Asn 280 | Leu | Met | Val | Phe | Met 285 | Asp | Tyr | Arg |
| Asp | Tyr 290 | Met | Gln | Gln | Lys | Leu 295 | Gln | Cys | Ser | Glu | Glu 300 | Glu | Tyr | Pro | Thr |
| Phe 305 | Leu | Tyr | Val | Met | Pro 310 | Met | Ser | Pro | Thr | Arg 315 | Leu | Phe | Phe | Glu | Glu 320 |
| Thr | Cys | Leu | Ala | Ser 325 | Lys | Asp | Ala | Met | Pro 330 | Phe | Asp | Leu | Leu | Lys 335 | Arg |
| Lys | Leu | Met | Ser 340 | Arg | Leu | Lys | Thr | Leu 345 | Gly | Ile | Gln | Val | Thr 350 | Lys | Ile |
| Tyr | Glu | Glu 355 | Glu | Trp | Ser | Tyr | Ile 360 | Pro | Val | Gly | Gly | Ser 365 | Leu | Pro | Asn |
| Thr | Glu 370 | Gln | Lys | Asn | Leu | Ala 375 | Phe | Gly | Ala | Ala | Ala 380 | Ser | Met | Val | His |
| Pro 385 | Ala | Thr | Gly | Tyr | Ser 390 | Val | Val | Arg | Ser | Leu 395 | Ser | Glu | Ala | Pro | Lys 400 |
| Tyr | Ala | Ser | Val | Ile 405 | Ala | Lys | Ile | Leu | Lys 410 | Gln | Asp | Asn | Ser | Ala 415 | Tyr |
| Val | Val | Ser | Gly 420 | Gln | Ser | Ser | Ala | Val 425 | Asn | Ile | Ser | Met | Gln 430 | Ala | Trp |
| Ser | Ser | Leu 435 | Trp | Pro | Lys | Glu | Arg 440 | Lys | Arg | Gln | Arg | Ala 445 | Phe | Phe | Leu |
| Phe | Gly 450 | Leu | Glu | Leu | Ile | Val 455 | Gln | Leu | Asp | Ile | Glu 460 | Ala | Thr | Arg | Thr |
| Phe 465 | Phe | Arg | Thr | Phe | Phe 470 | Arg | Leu | Pro | Thr | Trp 475 | Met | Trp | Trp | .Gly | Phe 480 |
| Leu | Gly | Ser | Ser | Leu | Ser | Ser | Phe | Asp | Leu | Val | Leu | Phe | Ser | Met | Tyr |

485

490

495

Met Phe Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His 500 505 510

Leu Leu Ser Asp Pro Ser Gly Ala Val Met Val Lys Ala Tyr Leu Glu 515 520 525

Arg

<210> 52

<211> 533

<212> PRT

<213> Lettuce

<400> 52

Met Glu Cys Phe Gly Ala Arg Asn Met Thr Ala Thr Met Ala Val Phe
1 5 10 15

Thr Cys Pro Arg Phe Thr Asp Cys Asn Ile Arg His Lys Phe Ser Leu 20 25 30

Leu Lys Gln Arg Arg Phe Thr Asn Leu Ser Ala Ser Ser Ser Leu Arg 35 40 45

Gln Ile Lys Cys Ser Ala Lys Ser Asp Arg Cys Val Val Asp Lys Gln
50 60

Gly Ile Ser Val Ala Asp Glu Glu Asp Tyr Val Lys Ala Gly Gly Ser 65 70 . 75 80

Glu Leu Phe Phe Val Gln Met Gln Arg Thr Lys Ser Met Glu Ser Gln 85 90 95

Ser Lys Leu Ser Glu Lys Leu Ala Gln Ile Pro Ile Gly Asn Cys Ile 100 105 110

Leu Asp Leu Val Val Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala 115 120 125

Ala Glu Ser Ala Lys Leu Gly Leu Asn Val Gly Leu Ile Gly Pro Asp 130 135 140

Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Gln Asp Glu Phe Ile Gly
145 150 155 160

Leu Gly Leu Glu Gly Cys Ile Glu His Ser Trp Lys Asp Thr Leu Val

Tyr Leu Asp Asp Ala Asp Pro Ile Arg Ile Gly Arg Ala Tyr Gly Arg 180 185 190

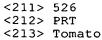
Val His Arg Asp Leu Leu His Glu Glu Leu Leu Arg Arg Cys Val Glu 195 200 205

Ser Gly Val Ser Tyr Leu Ser Ser Lys Val Glu Arg Ile Thr Glu Ala 210 215 220

Pro Asn Gly Tyr Ser Leu Ile Glu Cys Glu Gly Asn Ile Thr Ile Pro

225 230 235 Cys Arg Leu Ala Thr Val Ala Ser Gly Ala Ala Ser Gly Lys Phe Leu 245. Glu Tyr Glu Leu Gly Gly Pro Arg Val Cys Val Gln Thr Ala Tyr Gly Ile Glu Val Glu Val Glu Asn Asn Pro Tyr Asp Pro Asp Leu Met Val 280 Phe Met Asp Tyr Arg Asp Phe Ser Lys His Lys Pro Glu Ser Leu Glu. Ala Lys Tyr Pro Thr Phe Leu Tyr Val Met Ala Met Ser Pro Thr Lys Ile Phe Phe Glu Glu Thr Cys Leu Ala Ser Arg Glu Ala Met Pro Phe 330 Asn Leu Leu Lys Ser Lys Leu Met Ser Arg Leu Lys Ala Met Gly Ile Arg Ile Thr Arg Thr Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala 375 Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Asn Tyr Ala Ala Val Ile Ala Lys Ile Leu Arg Gln Asp Gln Ser Lys Glu Met Ile Ser Leu Gly Lys Tyr Thr Asn Ile Ser 425 Lys Gln Ala Trp Glu Thr Leu Trp Pro Leu Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu Phe Gly Leu Ser His Ile Val Leu Met Asp Leu Glu 455 Gly Thr Arg Thr Phe Phe Arg Thr Phe Phe Arg Leu Pro Lys Trp Met 470 Trp Trp Gly Phe Leu Gly Ser Ser Leu Ser Ser Thr Asp Leu Ile Ile 490 Phe Ala Leu Tyr Met Phe Val Ile Ala Pro His Ser Leu Arg Met Glu 505 Leu Val Arg His Leu Leu Ser Asp Pro Thr Gly Ala Thr Met Val Lys 520 Ala Tyr Leu Thr Ile 530

<210> 53



<400> 53

Met Glu Cys Val Gly Val Gln Asn Val Gly Ala Met Ala Val Leu Thr 1 5 10

Arg Pro Arg Leu Asn Arg Trp Ser Gly Glu Leu Cys Gln Glu Lys 20 25 30

Ser Ile Phe Leu Ala Tyr Glu Gln Tyr Glu Ser Lys Cys Asn Ser Ser 35 40 45

Ser Gly Ser Asp Ser Cys Val Val Asp Lys Glu Asp Phe Ala Asp Glu 50 55 60

Glu Asp Tyr Ile Lys Ala Gly Gly Ser Gln Leu Val Phe Val Gln Met 65 70 75 80

Gln Gln Lys Lys Asp Met Asp Gln Gln Ser Lys Leu Ser Asp Glu Leu 85 90 95

Arg Gln Ile Ser Ala Gly Gln Thr Val Leu Asp Leu Val Val Ile Gly
100 105 110

Cys Gly Pro Ala Gly Leu Ala Leu Ala Ala Glu Ser Ala Lys Leu Gly 115 120 125

Leu Asn Val Gly Leu Val Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr 130 135 140

Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Gln Ala Cys Ile 145 150 155 160

Glu His Val Trp Arg Asp Thr Ile Val Tyr Leu Asp Asp Asp Glu Pro 165 170 175

Ile Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Phe Leu His 180 185 190

Glu Glu Leu Leu Lys Arg Cys Val Glu Ala Gly Val Leu Tyr Leu Asn 195 200 205

Ser Lys Val Asp Arg Ile Val Glu Ala Thr Asn Gly Gln Ser Leu Val 210 215 220

Glu Cys Glu Gly Asp Val Val Ile Pro Cys Arg Phe Val Thr Val Ala 225 230 235 240

Ser Gly Ala Ala Ser Gly Lys Phe Leu Gln Tyr Glu Leu Gly Ser Pro 245 250 255

Arg Val Ser Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Asp Asn 260 265 270

Asn Pro Phe Asp Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp Tyr 275 280 285

Leu Arg His Asp Ala Gln Ser Leu Glu Ala Lys Tyr Pro Thr Phe Leu 290 295 300



| Tyr 305 | Ala | Met | Pro | Met | Ser 310 | Pro | Thr | Arg | Val | Phe 315 | Phe | Glu | Glu | Thr | Cys 320 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-------------|------------|------------|------------|------------|
| Leu | Ala | Ser | Lys | Asp 325 | Ala | Met | Pro | Phe | Asp 330 | Leu | Leu | Lys | Lys | Lys 335 | Leu |
| Met | Leu | Arg | Leu 340 | Asn | Thr | Leu | Gly | Val 345 | Arg | Ile | Lys | Glu | Ile 350 | Tyr | Glu |
| Glu | Glu | Trp 355 | Ser | Tyr | Ile | Pro | Val 360 | Gly | Gly | Ser | Leu | Pro 365 | Asn | Thr | Glu |
| Gln | Lys 370 | Thr | Leu | Ala | Phe | Gly 375 | Ala | Ala | Ala | Ser | 'Met 380 | Val | His | Pro | Ala |
| Thr 385 | Gly | Tyr | Ser | Val | Val 390 | Arg | Ser | Leu | Ser | Glu 395 | Ala | Pro | Lys | Cys | Ala 400 |
| Ser | Val | Leu | Ala | Asn 405 | Ile | Leu | Arg | Gln | His 410 | Tyr | Ser | Lys | Asn | Met 415 | Leu |
| Thr | Ser | Ser | Ser 420 | Ile | Pro | Ser | Ile | Ser 425 | Thr | Gln | Ala | Trp | Asn 430 | Thr | Leu |
| Trp | Pro | Gln 435 | Glu | Arg | Lys | Arg | Gln 440 | Arg | Ser | Phe | Phe | Leu 445 | Phe | Gly | Leu |
| Ala | Leu 450 | Ile | Leu | Gln | Leu | Asp 455 | Ile | Glu | Gly | Ile | Arg 460 | Ser | Phe | Phe | Arg |
| Ala 465 | Phe | Phe | Arg | Val | Pro 470 | Lys | Trp | Met | Trp | Gln 475 | Gly | Phe | Leu | Gly | Ser 480 |
| Ser | Leu | Ser | Ser | Ala 485 | Asp | Leu | Met | Leu | Phe 490 | Ala | Phe | Tyr | Met | Phe 495 | Ile |
| Ile | Ala | Pro | Asn 500 | Asp | Met | Arg | Lys | Gly 505 | Leu | Ile | Arg | His | Leu 510 | Leu | Ser |
| Asp | Pro | Thr 515 | Gly | Ala | Thr | Leu | Ile 520 | Arg | Thr | Tyr | Leu | Thr 525 | Phe | | |

<210> 54

<211> 516

<212> PRT

<213> Tagetes erecta

<400> 54

Met Ser Met Arg Ala Gly His Met Thr Ala Thr Met Ala Ala Phe Thr $1 \hspace{1cm} 5 \hspace{1cm} 10 \hspace{1cm} 15$

Cys Pro Arg Phe Met Thr Ser Ile Arg Tyr Thr Lys Gln Ile Lys Cys 20 25 30

Asn Ala Ala Lys Ser Gln Leu Val Val Lys Gln Glu Ile Glu Glu Glu 35 40 45

Glu Asp Tyr Val Lys Ala Gly Gly Ser Glu Leu Leu Phe Val Gln Met 50 55 60

WO 99/63055



| | | - | | | | | | | | | | | | | |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|-------------|------------|------------|------------|------------|------------|------------|
| Gln 65 | Gln | Asn | Lys | Ser | Met 70 | Asp | Ala | Gln | Ser | Ser 75 | Leu | Ser | Gln | Lys | Leu 80 |
| Pro | Arg | Val | Pro | Ile 85 | Gly | Gly | Gly | Gly | Asp 90 | Ser | Asn | Cys | Ile | Leu 95 | Asp |
| Leu | Val | Val | Ile 100 | Gly | Cys | Gly | Pro | Ala 105 | Gly | Leu | Ala | Leu | Ala 110 | Gly | Glu |
| Ser | Ala | Lys 115 | Leu | Gly | Leu | Asn | Val 120 | Ala | Leu | Ile | Gly | Pro 125 | Asp | Leu | Pro |
| Phe | Thr 130 | Așn | Asn | Tyr | Gly | Val 135 | Trp | Glu | Asp | Glu | Phe 140 | Ile | Gly | Leu | Gly |
| Leu 145 | Glu | Gly | Cys | Ile | Glu 150 | His | Val | Trp | Arg | Asp 155 | Thr | Val | Val | Tyr | Leu 160 |
| Asp | Asp | Asn | Asp | Pro 165 | Ile | Leu | Ile | Gly | 'Arg 170 | Ala | Tyr | Gly | Arg | Val 175 | Ser |
| Arg | Asp | Leu | Leu 180 | His | Glu | Glu | Leu | Leu 185 | Thr | Arg | Cys | Met | Glu 190 | Ser | Gly |
| Val | Ser | Tyr 195 | Leu | Ser | Ser | Lys | Val 200 | Glu | Arg | Ile | Thr | Glu 205 | Ala | Pro | Asn |
| Gly | Leu 210 | | Leu | Ile | Glu | Cys 215 | Glu | Gly | Asn | Ile | Thr 220 | Ile | Pro | Cys | Arg |
| Leu 225 | Ala | Thr | Val | Ala | Ser 230 | Gly | Ala | Ala | Ser | Gly 235 | _ | Leu | Leu | Gln | Tyr 240 |
| Glu | Leu | Gly | Gly | Pro 245 | Arg | Val | Суѕ | Val | Gln 250 | Thr | Ala | Tyr | Gly | Ile 255 | Glu |
| Val | Glu | Val | Glu 260 | Ser | Ile | Pro | Tyr | Asp 265 | Pro | Ser | Leu | Met | Val 270 | Phe | Met |
| Asp | Tyr | Arg 275 | Asp | Tyr | Thr | Lys | His 280 | Lys | Ser | Gln | Ser | Leu 285 | Glu | Ala | Gln |
| Tyr | Pro 290 | Thr | Phe | Leu | Tyr | Val 295 | Met | Pro | Met | Ser | Pro 300 | Thr | Lys | Val | Phe |
| Phe 305 | Glu | Glu | Thr | Cys | Leu 310 | | Ser | Lys | Glu | Ala 315 | | Pro | Phe | Glu | Leu 320 |
| Leu | Lys | Thr | Lys | Leu 325 | | Ser | Arg | Leu | Lys 330 | | Met | Gly | Ile | Arg 335 | Ile |
| Thr | Lys | Thr | Tyr 340 | | Glu | Glu | Trp | Ser 345 | | Ile | Pro | Val | Gly 350 | _ | Ser |
| Leu | Pro | Asn 355 | | Glu | Gln | Lys | Asn 360 | | Ala | Phe | Gly | Ala 365 | | Ala | Ser |
| Met | Val 370 | | Pro | Ala | Thr | Gly 375 | | Ser | Val | Val | Arg 380 | | Leu | Ser | Glu |
| n 1 | _ | | _ | | | | | | - | | - | | _ | | _ |

Ala Pro Asn Tyr Ala Ala Val·Ile Ala Lys Ile Leu Gly Lys Gly Asn

385 390 395 400 Ser Lys Gln Met Leu Asp His Gly Arg Tyr Thr Thr Asn Ile Ser Lys 405 Gln Ala Trp Glu Thr Leu Trp Pro Leu Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu Phe Gly Leu Ala Leu Ile Val Gln Met Asp Ile Glu Gly Thr Arg Thr Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp 455 Trp Gly Phe Leu Gly Ser Ser Leu Ser Ser Thr Asp Leu Ile Ile Phe Ala Phe Tyr Met Phe Ile Ile Ala Pro His Ser Leu Arg Met Gly Leu Val Arg His Leu Leu Ser Asp Pro Thr Gly Gly Thr Met Leu Lys Ala Tyr Leu Thr Ile 515 <210> 55 <211> 501 <212> PRT <213> Arabidopsis thaliana <400> 55 Met Asp Thr Leu Leu Lys Thr Pro Asn Lys Leu Asp Phe Phe Ile Pro Gln Phe His Gly Phe Glu Arg Leu Cys Ser Asn Asn Pro Tyr His Ser Arg Val Arg Leu Gly Val Lys Lys Arg Ala Ile Lys Ile Val Ser Ser Val Val Ser Gly Ser Ala Ala Leu Leu Asp Leu Val Pro Glu Thr Lys Lys Glu Asn Leu Asp Phe Glu Leu Pro Leu Tyr Asp Thr Ser Lys Ser 65 70 75 80 Gln Val Val Asp Leu Ala Ile Val Gly Gly Pro Ala Gly Leu Ala

Val Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile Asp 100 105 110

Pro Ser Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val Asp

Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Thr Thr Trp Ser

Gly Ala Val Val Tyr Val Asp Glu Gly Val Lys Lys Asp Leu Ser Arg



| | | - | | | | | | | | | • | | | | |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 145 | | | | • | 150 | | | | | 155 | | | | | 160 |
| Pro | Tyŗ | Gly | Arg | Val 165 | Asn | Arg | Lys | Gln | Leu 170 | Lys | Ser | Lys | Met | Leu 175 | Gln |
| Lys | Cys | Ile | Thr 180 | Asn | Gly | Val | Lys | Phe 185 | His | Gln | Ser | Lys | Val 190 | Thr | Asn |
| Val | Val | His 195 | Glu | Glu | Ala | Asn | Ser 200 | Thr | Val | Val | Cys | Ser 205 | Asp | Gly | Val |
| Lys | Ile 210 | Gln | Ala | Ser | Val | Val 215 | Leu | Asp | Ala | Thr | Gly 220 | Phe | Ser | Arg | Cys |
| Leu 225 | Val | Gln | Tyr | Asp | Lys 230 | Pro | Tyr | Asn | Pro | Gly 235 | Tyr | Gln | Val | Ala | Tyr 240 |
| Gly | Ile | Val | Ala | Glu 245 | Val | Asp | Gly | His | Pro 250 | Phe | Asp | Val | Asp | Lys 255 | Met |
| Val | Phe | Met | Asp 260 | Trp | Arg | Asp | Lys | His 265 | Leu | Asp | Ser | Tyr | Pro 270 | Glu | Leu |
| Lys | Glu | Arg 275 | Asn | Ser | Lys | Ile | Pro 280 | Thr | Phe | Leu | Tyr | Ala 285 | Met | Pro | Phe |
| Ser | Ser 290 | Asn | Arg | Ile | Phe | Leu 295 | Glu | Glu | Thr | Ser | Leu 300 | Val | Ala | Arg | Pro |
| Gly 305 | Leu | Arg | Met | Glu | Asp 310 | Ile | Gln | Glu | Arg | Met 315 | Ala | Ala | Arg | Leu | Lys 320 |
| His | Leu | Gly | Ile | Asn 325 | Val | Lys | Arg | Ile | Glu 330 | Glű | Asp | Glu | Arg | Cys 335 | Val |
| Ile | Pro | Met | Gly 340 | Gly | Pro | Leu | Pro | Val 345 | Leu | Pro | Gln | Arg | Val 350 | Val | Gly |
| Ile | Gly | Gly 355 | Thr | Ala | Gly | Met | Val 360 | His | Pro | Ser | Thr | Gly 365 | Tyr | Met | Val |
| Ala | Arg 370 | Thr | Leu | Ala | Ala | Ala 375 | Pro | Ile | Val | Ala | Asn 380 | Ala | Ile | Val | Arg |
| Tyr 385 | Leu | Gly | Ser | Pro | Ser 390 | Ser | Asn | Ser | Leu | Arg 395 | Gly | Asp | Gln | Leu | Ser 400 |
| Ala | Glu | Val | Trp | Arg 405 | | Leu | Trp | Pro | Ile 410 | Glu | Arg | Arg | Arg | Gln 415 | Arg |
| Glu | Phe | Phe | Cys 420 | Phe | Gly | Met | Asp | Ile 425 | Leu | Leu | Lys | Leu | Asp 430 | Leu | Asp |
| Ala | Thr | Arg 435 | Arg | Phe | Phe | Asp | Ala 440 | Phe | Phe | Asp | Leu | Gln 445 | Pro | His | Tyr |
| Trp | His 450 | Gly | Phe | Leu | Ser | Ser 455 | Arg | Leu | Phe | Leu | Pro 460 | Glu | Leu | Leu | Val |
| Phe 465 | Gly | Leu | Ser | Leu | Phe 470 | Ser | His | Ala | Ser | Asn 475 | Thr | Ser | Arg | Leu | Glu 480 |



Ile Met Thr Lys Gly Thr Val Pro Leu Ala Lys Met Ile Asn Asn Leu 485 490 495

Val Gln Asp Arg Asp 500

<210> 56

<211> 502

<212> PRT

<213> Adonis palaestina

<400> 56

Met Asp Thr Leu Leu Arg Thr His Asn Lys Leu Glu Leu Leu Pro Thr 1 5 10 15

Leu His Gly Phe Ala Glu Lys Gln His Leu Val Ser Thr Ser Lys Leu 20 25 30

Gln Asn Gln Val Phe Arg Ile Ala Ser Arg Asn Ile His Pro Cys Arg 35 40 45

Asn Gly Thr Val Lys Ala Arg Gly Ser Ala Leu Leu Glu Leu Val Pro 50 55 60

Glu Thr Lys Lys Glu Asn Leu Glu Phe Asp Leu Pro Ala Tyr Asp Pro 65 70 75 80

Ser Arg Gly Ile Val Val Asp Leu Ala Val Val Gly Gly Pro Ala 85 90 95

Gly Leu Ala Ile Ala Gln Gln Val Ser Glu Ala Gly Leu Leu Val Cys 100 105 110

Ser Ile Asp Pro Ser Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val 115 120 125

Trp Val Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Thr 130 135 140

Thr Trp Ser Gly Ala Val Val Tyr Thr Asp Asp Asn Ser Lys Lys Tyr 145 150 155 160

Leu Asp Arg Pro Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys 165 170 175

Met Leu Gln Lys Cys Val Thr Asn Gly Val Lys Phe His Gln Ala Lys 180 185 190

Val Ile Lys Val Ile His Glu Glu Ser Lys Ser Leu Leu Ile Cys Asn 195 200 205

Asp Gly Ile Thr Ile Asn Ala Thr Val Val Leu Asp Ala Thr Gly Phe 210 215 220

Ser Arg Cys Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln 225 230 235 240

Val Ala Tyr Gly Ile Met Ala Glu Val Glu Glu His Pro Phe Asp Leu 245 250 255



Asp Lys Met Leu Phe Met Asp Trp Arg Asp Ser His Leu Asn Glu Lys 260 265 270

Leu Glu Leu Lys Asp Lys Asn Arg Lys Ile Pro Thr Phe Leu Tyr Ala

Leu Glu Leu Lys Asp Lys Asn Arg Lys Ile Pro Thr Phe Leu Tyr Ala 275 280 285

Met Pro Phe Ser Ser Thr Lys Ile Phe Leu Glu Glu Thr Ser Leu Val 290 295 300

Ala Arg Pro Gly Leu Arg Phe Glu Asp Ile Gln Glu Arg Met Val Ala 305 310 315 320

Arg Leu Lys His Leu Gly Ile Lys Val Lys Ser Ile Glu Glu Asp Glu 325 330 335

Arg Cys Val Ile Pro Met Gly Gly Pro Leu Pro Val Leu Pro Gln Arg 340 345 350

Val Val Gly Ile Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly 355 360 365

Tyr Met Val Ala Arg Thr Leu Ala Ala Ala Pro Val Val Ala Lys Ser 370 375 380

Ile Val Gln Tyr Leu Gly Ser Asp Arg Ser Leu Ser Gly Asn Glu Leu 385 390 395 400

Ser Ala Glu Val Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Gln 405 410 415

Arg Glu Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu 420 425 430

Gln Gly Thr Arg Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro His 435 440 445

Tyr Trp His Gly Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Leu 450 460

Phe Phe Gly Leu Ser Leu Phe Ser His Ala Ser Asn Ala Ser Arg Ile 465 470 475 480

Glu Ile Met Ala Lys Gly Thr Val Pro Leu Val Asn Met Met Asn Asn 485 490 495

Leu Ile Gln Asp Thr Asp 500

<210> 57

<211> 498

<212> PRT

<213> Pepper

<400> 57

Met Asp Thr Leu Leu Arg Thr Pro Asn Asn Leu Glu Phe Leu His Gly
1 10 15

Phe Gly Val Lys Val Ser Ala Phe Ser Ser Val Lys Ser Gln Lys Phe 20 25 30

| | | - | | | | | | | | | | | | | |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Gly | Ala | Lys 35 | Lys | Phe | Cys | Glu | Gly 40 | Leu | Gly | Ser | Arg | Ser 45 | Val | Cys | Val |
| Lys | Ala 50 | Ser | Ser | Ser | Ala | Leu 55 | Leu | Glu | Leu | Val | Pro 60 | Glu | Thr | Lys | Lys |
| Glu 65 | Asn | Leu | Asp | Phe | Glu 70 | Leu | Pro | Met | Tyr | Asp 75 | Pro | Ser | Lys | Gly | Val 80 |
| Val | Val | Asp | Leu | Ala 85 | Val | Val | Gly | Gly | Gly 90 | Pro | Ala | Gly | Leu | Ala 95 | Val |
| Ala | Gln | Gln | Val 100 | Ser | Glu | Ala | Gly | Leu 105 | Ser | Val | Cys | Ser | Ile 110 | Asp | Pro |
| Asn | Pro | Lys 115 | Leu | Ile | Trp | Pro | Asn 120 | Asn | Tyr | Gly | Val | Trp 125 | Val | Asp | Glu |
| Phe | Glu 130 | Ala | Met | Asp | Leu | Leu 135 | Asp | Cys | Leu | Asp | Ala 140 | Thr | Trp | Ser | Gly |
| Ala 145 | Ala | Val | Tyr | Ile | Asp 150 | Asp | Lys | Thr | Thr | Lys 155 | Asp | Leu | Asn | Arg | Pro 160 |
| Tyr | Gly | Arg | Val | Asn 165 | Arg | Lys | Gln | Leu | Lys 170 | Ser | Lys | Met | Met | Gln 175 | Lys |
| Cys | Ile | Leu | Asn 180 | Gly | Val | Lys | Phe | His 185 | Gln | Ala | Lys | Val | Ile 190 | Lys | Val |
| Ile | His | Glu 195 | Glu | Ser | Lys | Ser | Met 200 | Leu | Ile | Cys | Asn | Asp 205 | Gly | Ile | Thr |
| Ile | Gln 210 | Ala | Thr | Val | Val | Leu 215 | Asp | Ala | Thr | Gly | Phe 220 | Ser | Arg | Ser | Leu |
| Val 225 | Gln | Tyr | Asp | Lys | Pro 230 | Tyr | Asn | Pro | Gly | Tyr 235 | Gln | Val | Ala | Tyr | Gly 240 |
| Ile | Leu | Ala | Glu | Val 245 | Glu | Glu | His | Pro | Phe 250 | Asp | Val | Asn | Lys | Met 255 | Val |
| Phe | Met | Asp | Trp 260 | Arg | Asp | Ser | His | Leu 265 | Lys | Asn | Asn | Val | Glu 270 | Leu | Lys |
| Glu | Arg | Asn 275 | Ser | Arg | Ile | Pro | Thr 280 | Phe | Leu | Tyr | Ala | Met 285 | Pro | Phe | Ser |
| Ser | Asn 290 | Arg | Ile | Phe | Leu | Glu 295 | Glu | Thr | Ser | Leu | Val 300 | Ala | Arg | Pro | Gly |
| Leu 305 | Gly | Met | Asp | Asp | Ile 310 | Gln | Glu | Arg | Met | Val 315 | Ala | Arg | Leu | Ser | His 320 |
| Leu | Gly | Ile | Lys | Val 325 | Lys | Ser | Ile | Glu | Glu 330 | Asp | Glu | His | Cys | Val 335 | Ile |
| Pro | Met | Gly | Gly 340 | Pro | Leu | Pro | Val | Leu 345 | Pro | Gln | Arg | Val | Val 350 | Gly | Ile |
| Gly | Gly | Thr | Ala | Gly | Met | Val | His | Pro | Ser | Thr | Gly | Tyr | Met | Val | Ala |
| | | | | | | | | | | | | | | | |

- 1

355 360 365

Arg Thr Leu Ala Ala Ala Pro Val Val Ala Asn Ala Ile Ile Gln Tyr 370 380

Leu Ser Ser Glu Arg Ser His Ser Gly Asp Glu Leu Ser Ala Ala Val 385 390 395 400

Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Gln Arg Glu Phe Phe 405 410 415

Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu Pro Ala Thr Arg 420 425 430

Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp His Gly 435 440 445

Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Ile Val Phe Gly Leu 450 455 460

Ser Leu Phe Ser His Ala Ser Asn Thr Ser Arg Leu Glu Ile Met Thr 465 470 475 480

Lys Gly Thr Leu Pro Leu Val His Met Ile Asn Asn Leu Leu Gln Asp 485 490 495

Lys Glu

<210> 58

<211> 500

<212> PRT

<213> Tomato

<400> 58

Met Asp Thr Leu Leu Lys Thr Pro Asn Asn Leu Glu Phe Leu Asn Pro 1 5 10 15

His His Gly Phe Ala Val Lys Ala Ser Thr Phe Arg Ser Glu Lys His 20 25 30

His Asn Phe Gly Ser Arg Lys Phe Cys Glu Thr Leu Gly Arg Ser Val 35 40 45

Cys Val Lys Gly Ser Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr 50 55 60

Lys Lys Glu Asn Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys 65 70 75 80

Gly Val Val Asp Leu Ala Val Val Gly Gly Pro Ala Gly Leu 85 90 95

Ala Val Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile 100 105 110

Asp Pro Asn Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val 115 120 125

Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp



| | 130 | | | | | 135 | | | | | 140 | | | | 7 _: |
|------------|------------|------------|------------|------------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Ser 145 | Gly | Ala | Ala | | Tyr .150 | Ile | Asp | Asp | Asn | Thr 155 | Ala | Lys | Asp | Leu | His 160 |
| Arg | Pro | Tyr | Gly | Arg 165 | Val | Asn | Arg | Lys | Gln 170 | Leu | Lys | Ser | Lys | Met 175 | Met |
| Gln | Lys | Cys | Ile 180 | Met | Asn | Gly | Val | Lys 185 | Phe | His | Gln | Ala | Lys 190 | Val | Ile |
| Lys | Val | Ile 195 | His | Glu | Glu | Ser | Lys 200 | Ser | Met | Leu | Ile | Cys 205 | Asn | Asp | Gly |
| Ile | Thr 210 | Ile | Gln | Ala | Thr | Val 215 | Val | Leu | Asp | Ala | Thr 220 | Gly | Phe | Ser | Arg |
| Ser 225 | Leu | Val | Gln | Tyr | Asp 230 | Lys | Pro | Tyr | Asn | Pro 235 | Gly | Tyr | Gln | Val | Ala 240 |
| Tyr | Gly | Ile | Leu | Ala 245 | Glu | Val | Glu | Glu | His 250 | Pro | Phe | Asp | Val | Asn 255 | Lys |
| Met | Val | Phe | Met 260 | Asp | Trp | Arg | Asp | Ser 265 | His | Leu | Lys | Asn | Asn 270 | Thr | Asp |
| Leu | Lys | Glu 275 | Arg | Asn | Ser | Arg | Ile 280 | Pro | Thr | Phe | Leu | Tyr 285 | Ala | Met | Pro |
| Phe | Ser 290 | Ser | Asn | Arg | Ile | Phe 295 | Leu | Glu | Glu | Thr | Ser 300 | Leu | Val | Ala | Arg |
| Pro 305 | Gly | Leu | Arg | Ile | Asp 310 | Asp | Ile | Gln | Glu | Arg 315 | Met | Val | Ala | Arg | Leu 320 |
| Asn | His | Leu | Gly | 11e 325 | Lys | Val | Lys | Ser | 11e 330 | Glu | Glu | Asp | Glu | His 335 | Cys |
| Leu | Ile | Pro | Met 340 | Gly | Gly | Pro | Leu | Pro 345 | Val | Leu | Pro | Gln | Arg 350 | Val | Val |
| Gly | Ile | Gly 355 | Gly | Thr | Ala | Gly | Met 360 | Val | His | Þго | Ser | Thr 365 | Gly | Tyr | Met |
| Val | Ala 370 | Arg | Thr | Leu | Ala | Ala 375 | Ala | Pro | Val | Val | Ala 380 | Asn | Ala | Ile | Ile |
| Gln 385 | Tyr | Leu | Gly | Ser | Glu 390 | Ärg | Ser | His | Ser | Gly 395 | Asn | Glu | Leu | Ser | Thr 400 |
| Ala | Val | Trp | Lys | Asp 405 | Leu | Trp | Pro | Ile | Glu 410 | Arg | Arg | Arg | Gln | Arg 415 | Glu |
| Phe | Phe | Cys | Phe 420 | Gly | Met | Asp | Ile | Leu 425 | Leu | Lys | Leu | Asp | Leu 430 | Pro | Ala |
| Thr | Arg | Arg 435 | Phe | Phe | Asp | Ala | Phe 440 | Phe | Asp | Leu | Glu | Pro 445 | Arg | Tyr | Trp |
| His | Gly 450 | Phe | Leu | Ser | Ser | Arg 455 | | Phe | Leu | Pro | Glu 460 | Leu | Ile | Val | Phe |
| | | | | | | | | | | | | | | | |



Gly Leu Ser Leu Phe Ser His Ala Ser Asn Thr Ser Arg Phe Glu Ile 465 470 475 480

Met Thr Lys Gly Thr Val Pro Leu Val Asn Met Ile Asn Asn Leu Leu 485 490 495

Gln Asp Lys Glu 500

<210> 59

<211> 500

<212> PRT

<213> Tobacco

<400> 59

Met Asp Thr Leu Leu Lys Thr Pro Asn Lys Leu Glu Phe Leu His Pro
1 5 10 15

Val His Gly Phe Ser Val Lys Ala Ser Ser Phe Asn Ser Val Lys Pro 20 25 30

His Lys Phe Gly Ser Arg Lys Ile Cys Glu Asn Trp Gly Lys Gly Val 35 40 45

Cys Val Lys Ala Lys Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr 50 55 60

Lys Lys Glu Asn Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys 65 70 75 80

Gly Leu Val Val Asp Leu Ala Val Val Gly Gly Pro Ala Gly Leu 85 90 95

Ala Val Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Val Ser Ile 100 105 110

Asp Pro Ser Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val 115 120 125

Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp 130 140

Ser Gly Thr Val Val Tyr Ile Asp Asp Asn Thr Thr Lys Asp Leu Asp 145 150 155 160

Arg Pro Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys Met Met
165 170 175

Gln Lys Cys Ile Leu Asn Gly Val Lys Phe His His Ala Lys Val Ile 180 185 190

Lys Val Ile His Glu Glu Ala Lys Ser Met Leu Ile Cys Asn Asp Gly 195 200 205

Val Thr Ile Gln Ala Thr Val Val Leu Asp Ala Thr Gly Phe Ser Arg 210 215 220

Cys Leu Val Gln Tyr Asp Lys Pro Tyr Lys Pro Gly Tyr Gln Val Ala 225 230 235 240



| Tyr | Gly | Ile | Leu | Ala 245 | Glu | Val | Glu | Glu | His 250 | Pro | Phe | Asp | Thr | Ser 255 | Lys |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-------------------------|------------|
| Met | Val | Leu | Met 260 | Asp | Trp | Arg | Asp | Ser 265 | His | Leu | Gly | Asn | Asn 270 | Met | Glu |
| Leu | Lys | Glu 275 | Arg | Asn | Arg | Lys | Val 280 | Pro | Thr | Phe | Leu | Tyr 285 | Ala | Met | Pro |
| Phe | Ser 290 | Ser | Asn | Lys | Ile | Phe 295 | Leu | Glu | Glu | Thr | Ser 300 | Leu | Val | Ala | Arg |
| Pro 305 | Gly | Leu | Arg | Met | Asp 310 | Asp | Ile | Gln | Glu | Arg 315 | Met | Val | Ala | Arg | Leu 320 |
| Asn | His | Leu | Gly | Ile 325 | Lys | Val | Lys | Ser | Ile 330 | Glu | Glu | Asp | Glu | His 335 | Cys |
| Val | Ile | Pro | Met 340 | Gly | Gly | Ser | Leu | Pro 345 | Val | Ile | Pro | Gln | Arg 350 | Val | Val |
| Gly | Thr | Gly 355 | Gly | Thr | Ala | Gly | Leu 360 | Val | His | Pro | Ser | Thr 365 | Gly | Tyr | Met |
| Val | Ala 370 | Arg | Thr | Leu | Ala | Ala 375 | Ala | Pro | Val | Val | Ala 380 | Asn | Ala | Ile | Ile |
| His 385 | Tyr | Leu | Gly | Ser | Glu 390 | Lys | Asp | Leu | Leu | Gly 395 | Asn | Glu | Leu | Ser | Ala 400 |
| Ala | Val | Trp | Lys | Asp 405 | Leu | Trp | Pro | Ile | Glu 410 | Arg | Arg | Arg | Gln | Arg 415 _. | |
| Phe | Phe | Cys | Phe 420 | Gly | Met | Asp | Ile | Leu 425 | Leu | Lys | Leu | Asp | Leu 430 | Pro | Ala |
| Thr | Arg | Arg 435 | Phe | Phe | Asp | Ala | Phe 440 | Phe | Asp | Leu | Glu | Pro 445 | Arg | Tyr | Trp |
| His | Gly 450 | Phe | Leu | Ser | Ser | Arg 455 | Leu | Tyr | Leu | Pro | Glu 460 | Leu | Ile | Phe | Phe |
| Gly 465 | Leu | Ser | Leu | Phe | Ser 470 | Arg | Ala | Ser | Asn | Thr 475 | Ser | Arg | Ile | Glu | Ile 480 |
| Met | Thr | Lys | Gly | Thr 485 | Leu | Pro | Leu | Val | Asn 490 | Met | Ile | Asn | Asn | Leu 495 | Leu |

Gln Asp Thr Glu 500

<210> 60 <211> 511 <212> PRT <213> Tagetes erecta

<400> 60

Met Asp Thr Phe Leu Arg Thr Tyr Asn Ser Phe Glu Phe Val His Pro 1 5 10 15



| | | | | 3.5 | 7 | | | | | | | | | | |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Ser | Asn | Lys | Phe 20 | Àla | Gly | Asn | Leu | Asn 25 | Asn | Leu | Asn | Gln | Leu 30 | Asn | Gln |
| Ser | Lys | Ser .35 | Gln | Phe | Gln | Asp | Phe 40 | .Arg | Phe | Gly | Pro | Lys 45 | Lys | Ser | Gln |
| Phe | Lys 50 | Leu | Gly | Gln | Lys | Tyr 55 | Cys | Val | Lys | Ala | Ser 60 | Ser | Ser | Ala | Leu |
| Leu 65 | Glu | Leu | Val | Pro | Glu 70 | Ile | Lys | Lys | Glu | Asn 75 | Leu | Asp | Phe | Asp | Leu 80 |
| Pro | Met | Tyr | Asp | Pro 85 | Ser | Arg | Asn | Val | Val 90 | Val | Asp | Leu | Val | Val 95 | Val |
| Gly | Gly | Gly | Pro 100 | Ser | Gly | Leu | Ala | Val 105 | Ala | Gln | Gln | Val | Ser 110 | Glu | Ala |
| Glý | Leu | Thr 115 | Val | Cys | Ser | Ile | Asp 120 | Pro | Ser | Pro | Lys | Leu 125 | Ile | Trp | Pro |
| Asn | Asn 130 | Tyr | Gly | Val | Trp | Val 135 | Asp | Glu | Phe | Glu | Ala 140 | Met | Asp | Leu | Leu |
| Asp 145 | Cys | Leu | Asp | Thr | Thr 150 | Trp | Ser | Ser | Ala | Val 155 | Val | Tyr | Ile | Asp | Glu 160 |
| Lys | Ser | Thr | Lys | Ser 165 | Leu | Asn | Arg | Pro | Tyr 170 | Ala | Arg | Val | Asn | Arg 175 | Lys |
| Gln | Leu | Lys | Thr 180 | Lys | Met | Leu | Glņ | Lys 185 | Cys | Ile | Ala | Asn | Gly 190 | Val | Lys |
| Phe | His | Gln 195 | Ala | Lys | Val | Ile | Lys 200 | Val | Ile | His | Glu | Glu 205 | Leu | Lys | Ser |
| Leu | Leu 210 | lle | Cys | Asn | Asp | Gly 215 | .Val | Thr | Ile | Gln | Ala 220 | Thr | Leu | Val | Leu |
| Asp 225 | Ala | Thr | Gly | Phe | Ser 230 | Arg | ·Ser | Leu | Val | Gln 235 | Tyr | Asp | Lys | Pro | Tyr 240 |
| Asn | Pro | Gly | Tyr | Gln 245 | Val | Ala | Tyr | Gly | 11e 250 | Leu | Ala | Glu | Val | Glu 255 | Glu |
| His | Pro | Phe | Asp 260 | Val | Asp | Lys | Met | Leu 265 | Phe | Met | Asp | Trp | Arg 270 | Asp | Ser |
| His | Leu | Asp 275 | Gln | Asn | Leu | Glu | Ile 280 | Lys | Ala | Arg | Asn | Ser 285 | Arg | Ile | Pro |
| Thr | Phe 290 | Leu | Tyr | Ala | Met | Pro 295 | Phe | Ser | Ser | Thr | Arg 300 | Ile | Phe | Leu | Glu |
| Glu 305 | Thr | Ser | Leu | Val | Ala 310 | Arg | Pro | Gly | Leu | Lys 315 | Met | Glu | Asp | Ile | Gln 320 |
| Glu | Arg | Met | Ala | Tyr 325 | Arg | Leu | Lys | His | Leu 330 | Gly | Ile | Lys | Val | Lys 335 | |
| Ile | Glu | Glu | Asp | Glu | Arg | Cys | Val | Ile | Pro | Met | Gly | Gly | Pro | Leu | Pro |

340 345

390

350

Val Leu Pro Gln Arg Val Leu Gly Ile Gly Gly Thr Ala Gly Met Val 355 360 365

His Pro Ser Thr Gly Tyr Met Val Ala Arg Thr Leu Ala Ala Pro

Ile Val Ala Lys Ser Ile Ile Arg Tyr Leu Asn Asn Glu Lys Ser Met

Val Ala Asp Val Thr Gly Asp Asp Leu Ala Ala Gly Ile Trp Arg Glu .

395

Leu Trp Pro Ile Glu Arg Arg Gln Arg Glu Phe Phe Cys Phe Gly 420 425 430

Met Asp Ile Leu Lys Leu Asp Leu Glu Gly Thr Arg Arg Phe Phe 435 440 445

Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp His Gly Phe Leu Ser 450 460

Ser Arg Leu Phe Leu Pro Glu Leu Val Thr Phe Gly Leu Ser Leu Phe 465 470 475 480

Gly His Ala Ser Asn Thr Cys Arg Val Glu Ile Met Ala Lys Gly Thr 485 490 495

Leu Pro Leu Ala Thr Met Ile Gly Asn Leu Val Arg Asp Arg Glu 500 505 510

<210> 61

<211> 503

<212> PRT

<213> Daffodil

<400> 61

Met Asp Thr Leu Leu Arg Thr His Asn Arg Leu Glu Leu Leu Tyr Pro 1 5 10 15

Leu His Glu Leu Ala Lys Arg His Phe Leu Ser Pro Ser Pro Asn Pro 20 25 30

Gln Asn Pro Asn Phe Lys Phe Phe Ser Arg Lys Pro Tyr Gln Lys Lys 35 40 45

Cys Arg Asn Gly Tyr Ile Gly Val Ser Ser Asn Gln Leu Leu Asp Leu
50 60

Val Pro Glu Ile Lys Lys Glu His Leu Glu Phe Asp Leu Pro Leu Tyr 65 70 75 80

Asp Pro Ser Lys Ala Leu Thr Leu Asp Leu Ala Val Val Gly Gly 85 90 95

Pro Leu Ala Arg Ser Cys Ser Thr Ser Leu Gly Gly Gly Leu Ser Val

Val Ser Ile Asp Pro Asn Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly

بن أجر

 Val
 Trp 130
 Val Asp Glu Phe 135
 Glu Asp Met Asp Leu Leu Asp 140
 Leu Asp Cys Leu Asp 140

 Ala Thr Trp Ser Gly Ala 150
 Tyr Val Asp 155
 Asp Arg Ser Thr Lys 160

Asn Leu Ser Arg Pro Tyr Ala Arg Val Asn Arg Lys Asn Leu Lys Ser

Lys Met Met Lys Lys Cys Val Ser Asn Gly Val Arg Phe His Gln Ala 180 185 190

Thr Val Val Lys Ala Met His Glu Glu Glu Lys Ser Tyr Leu Ile Cys 195 200 205

Ser Asp Gly Val Thr Ile Asp Ala Arg Val Val Leu Asp Ala Thr Gly 210 215 220

Phe Ser Arg Cys Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr 225 230 235 240

Gln Val Ala Tyr Gly Ile Leu Ala Glu Val Glu Glu His Pro Phe Asp 245 250 255

Val Asp Lys Met Val Phe Met Asp Trp Arg Asp Ser His Leu Asn Gly
260 265 270

Lys Ala Glu Leu Asn Glu Arg Asn Ala Lys Ile Pro Thr Phe Leu Tyr 275 280 285

Ala Met Pro Phe Ser Ser Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu 290 295 300

Val Ala Arg Pro Gly Leu Lys Met Glu Asp Ile Gln Glu Arg Met Val 305 310 315 320

Ala Arg Leu Asn His Leu Gly Ile Arg Ile Lys Ser Ile Glu Glu Asp 325 330 335

Glu Arg Cys Val Ile Pro Met Gly Gly Pro Leu Pro Val Ile Pro Gln 340 345 350

Arg Val Val Gly Ile Gly Gly Thr Ala Gly Met Val His Pro Ser Thr 355 360 365

Gly Tyr Met Val Ala Arg Thr Leu Ala Ala Ala Pro Ile Val Ala Asn 370 375 380

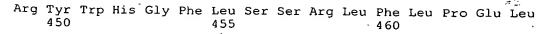
Ser Ile Val Gln Tyr Leu Val Ser Asp Ser Gly Leu Ser Gly Asn Asp 385 390 395 400

Leu Ser Ala Asp Val Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Arg 405 410 415

Gln Arg Glu Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp 420 425 430

Leu Glu Gly Thr Arg Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro 435 440 445





Val Pro Phe Gly Leu Ser Leu Phe Ser His Ala Ser Asn Thr Cys Lys 465 470 475 480

Leu Glu Ile Met Ala Lys Gly Thr Leu Pro Leu Val Asn Met Ile Asn 485 490 495

Asn Leu Val Gln Asp Arg Asp 500

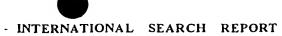


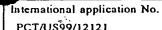
INTERNATIONAL SEARCH REPORT

International application No.

| | | | PC1/US99/1212 | :1 |
|--|---|----------------------------------|--|---|
| IPC(6) :Please Se US CL :435/189, | TION OF SUBJECT MATTER te Extra Sheet. 193, 233, 252.3, 320.1, 325; 536/23.2 tional Patent Classification (IPC) or to both r | national classification | and IPC | - |
| B. FIELDS SEAF | RCHED | | | ······································ |
| Minimum documenta | tion searched (classification system followed | by classification sym | ibols) | |
| | 193, 233, 252.3, 320.1, 325; 536/23.2 | , | · | |
| Documentation search | ed other than minimum documentation to the | extent that such docur | ments are included | in the fields searched |
| Electronic data base of Please See Extra Si | consulted during the international search (na | me of data base and, | where practicable, | search terms used) . |
| C. DOCUMENTS | S CONSIDERED TO BE RELEVANT | 1 | | |
| Category* Cital | tion of document, with indication, where app | propriate, of the releva | int passages | Relevant to claim No. |
| X WO 9 PARK No:1. | 97/36998 A1 (UNIVERSITY OF Colors of Colors 1997, see entire d | F MARYLAND ocument, especia | COLLEGE ally SEQ ID | 1-8 |
| | | | | |
| | | | | |
| | | | | • |
| Further docum | ents are listed in the continuation of Box C. | See paten | t family annex. | |
| | nes of cited documents: | | | rnational filing date or priority |
| "A" document defir to be of particu | ing the general state of the art which is not considered plar relevance | | theory underlying the | |
| *L* document which | nt published on or after the international filing date th may throw doubts on priority claim(s) or which is | considered nov | | e claimed invention cannot be red to involve an inventive step |
| special reason | | | | claimed invention cannot be step when the document is |
| nieans | rring to an oral disclosure, use, exhibition or other | being obvious | one or more other suct to a person skilled in t | n documents, such combination he art |
| the priority dat | | | ber of the same patent | |
| 02 AUGUST :999 | mpletion of the international search | Date of mailing of th | | rch report |
| Name and mailing ad Commissioner of Pate Box PCT | nts and Trademarks | Authorized officer BRADLEY S. M. | AYHEW A | U Y. |
| Washington, D.C. 20 Facsimile No. (703 | 231 3) 305-3230 | Telephone No. (7 | 03) 308-0196 | YUL |

BNSDOCID: <WO__9963055A1_I_>





| <u>-</u> | | | PCT/U | S99/12121 | |
|---|--------------------------|------------------------|--------------------|------------------|------------|
| A. CLASSIFICATION OF S IPC (6): | UBJECT MATTER: | | | | - |
| C12N 1/21, 5/10, 9/10, 15/53 | 3, 15/54, 15/61, 15/63; | C12P 23/00; C12Q 1/ | 68 | | |
| B. FIELDS SEARCHED Electronic data bases consult | ed (Name of data base) | and where practicable | terms used): | | |
| Dialog and APS search terms: IPP, epsilon c isomerase | yclase, lycopene cyclase | e, isopentenyl pyropho | sphate isomerase a | nd isopentenyl d | iphosphate |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |

Form PCT:ISA 210 (extra sheet/July 1992)*

PCT





INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6:

C12N 1/21, 5/10, 9/10, 15/53, 15/54, 15/61, 15/63, C12P 23/00, C12Q 1/68 (11) International Publication Number:

WO 99/63055

A1

(43) International Publication Date:

9 December 1999 (09.12.99)

(21) International Application Number:

PCT/US99/12121

(22) International Filing Date:

2 June 1999 (02.06.99)

(30) Priority Data:

09/088,724 09/088,725

2 June 1998 (02.06.98) 2 June 1998 (02.06.98) LIS US

(71) Applicant (for all designated States except US): UNIVERSITY OF MARYLAND [US/US]; Office of Technology Liaison, 4312 Knox Road, College Park, MD 20742 (US).

(72) Inventors; and

- (75) Inventors/Applicants (for US only): CUNNINGHAM, Francis, X., Jr. [US/US]; 2727 Washington Avenue, Chevy Chase, MD 20815 (US). SUN, Zairen [US/US]; 3405 Tulane Drive #22, Hyattsville, MD 20783 (US).
- (74) Agents: GOLDHUSH, Douglas, H. et al.; Nikaido, Marmelstein, Murray & Oram LLP, Suite 330 G Street Lobby, Metropolitan Square, 655 Fifteenth Street, N.W., Washington, DC 20005-5701 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report. With amended claims.

(54) Title: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND METHODS OF USE THEREOF

(57) Abstract

Nucleic acid sequences encoding ϵ -cyclase, isopentenyl pyrophosphate isomerase and β -carotene hydroxylase as well as vectors containing the same and hosts transformed with the vectors. Methods for controlling the ratio of various carotenoids in a host and for the production of novel carotenoid pigments. The present invention also provides a method for screening for eukaryotic genes encoding carotenoid biosynthesis, and for modifying the disclosed enzymes.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| AL | Albania | ES | Spain | LS | Lesotho | SI | Slovenia |
|----|--------------------------|----|---------------------|----|-----------------------|----|--------------------------|
| AM | Armenia | FI | Finland | LT | Lithuania | SK | Slovakia |
| AT | Austria | FR | France | LU | Luxembourg | SN | Senegal |
| AU | Australia | GA | Gabon | LV | Latvia | SZ | Swaziland |
| AZ | Azerbaijan | GB | United Kingdom | MC | Monaco | TD | Chad |
| BA | Bosnia and Herzegovina | GE | Georgia | MD | Republic of Moldova | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagascar | TJ | Tajikistan |
| BE | Belgium | GN | Guinea | MK | The former Yugoslav | TM | Turkmenistan |
| BF | Burkina Faso | GR | Greece | | Republic of Macedonia | TR | Turkey |
| BG | Bulgaria | HU | Hungary | ML | Mali | TT | Trinidad and Tobago |
| BJ | Benin | IE | Ireland | MN | Mongolia | UA | Ukraine |
| BR | Brazil | IL | Israel | MR | Mauritania | UG | Uganda |
| BY | Belarus | IS | Iceland | MW | Malawi | US | United States of America |
| CA | Canada | IT | Itały | MX | Mexico | UZ | Uzbekistan |
| CF | Central African Republic | JP | Japan | NE | Niger | VN | Viet Nam |
| CG | Congo | KE | Kenya | NL | Netherlands | YU | Yugoslavia |
| СН | Switzerland | KG | Kyrgyzstan | NO | Norway | zw | Zimbabwe |
| CI | Côte d'Ivoire | KP | Democratic People's | NZ | New Zealand | | |
| CM | Cameroon | | Republic of Korea | PL | Poland | | |
| CN | China | KR | Republic of Korea | PT | Portugal | | |
| CU | Cuba | KZ | Kazakstan | RO | Romania | | |
| CZ | Czech Republic | LC | Saint Lucia | RU | Russian Federation | | |
| DE | Germany | LI | Liechtenstein | SD | Sudan | | |
| DK | Denmark | LK | Sri Lanka | SE | Sweden | | |
| EE | Estonia | LR | Liberia | SG | Singapore | | |
| 1 | | | | | | | |

GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND METHODS OF USE THEREOF

BACKGROUND OF THE INVENTION

Field of the Invention

5

10

The present invention describes nucleic acid sequences for eukaryotic genes encoding ϵ lycopene ϵ -cyclase (also known as ϵ -cyclase and ϵ lycopene cyclase), isopentenyl pyrophosphate isomerase (IPP) and β -carotene hydroxylase as well as vectors containing the same and hosts transformed with said vectors. The present invention also provides methods for augmenting the accumulation of carotenoids, changing the composition of the carotenoids, and producing novel and rare carotenoids. The present invention provides methods for controlling the ratio or relative amounts of various carotenoids in a host. The invention also relates to modified lycopene ϵ -cyclase, IPP isomerase and β -carotene hydroxylase. Additionally, the present invention provides a method for screening for genes and cDNAs encoding enzymes of carotenoid biosynthesis and metabolism.

15

20

25

30

Background of the Invention

Carotenoid pigments with cyclic endgroups are essential components of the photosynthetic apparatus in oxygenic photosynthetic organisms (e.g., cyanobacteria, algae and plants; Goodwin, 1980). The symmetrical bicyclic yellow carotenoid pigment βcarotene (or, in rare cases, the asymmetrical bicyclic α-carotene) is intimately associated with the photosynthetic reaction centers and plays a vital role in protecting against potentially lethal photooxidative damage (Koyama, 1991). β-carotene and other carotenoids derived from it or from α-carotene also serve as light-harvesting pigments (Siefermann-Harms, 1987), are involved in the thermal dissipation of excess light energy captured by the lightharvesting antenna (Demmig-Adams & Adams, 1992), provide substrate for the biosynthesis of the plant growth regulator abscisic acid (Rock & Zeevaart, 1991; Parry & Horgan, 1991), and are precursors of vitamin A in human and animal diets (Krinsky, 1987). Plants also exploit carotenoids as coloring agents in flowers and fruits to attract pollinators and agents of seed dispersal (Goodwin, 1980). The color provided by carotenoids is also of agronomic value in a number of important crops. Carotenoids are currently harvested from a variety of organisms, including plants, algae, yeasts, cyanobacteria and bacteria, for use as pigments in food and feed.

WO 99/63055

The probable pathway for formation of cyclic carotenoids in plants, algae and cyanobacteria is illustrated in Figure 1. Two types of cyclic endgroups or rings are commonly found in higher plant carotenoids, these are referred to as the β (beta) and ϵ (epsilon) rings (Fig. 3). The precursor acyclic endgroup (no ring structure) is referred to as the Ψ (psi) endgroup. The β and ϵ endgroups differ only in the position of the double bond in the ring. Carotenoids with two β rings are ubiquitous, and those with one β and one ϵ ring are common, but carotenoids with two ϵ rings are uncommon. β -carotene (Fig. 1) has two β -endgroups and is a symmetrical compound that is the precursor of a number of other important plant carotenoids such as zeaxanthin and violaxanthin (Fig. 2).

10

15

20

5

Genes encoding enzymes of carotenoid biosynthesis have previously been isolated from a variety of sources including bacteria (Armstrong et al., 1989, Mol. Gen. Genet. 216, 254-268; Misawa et al., 1990, J. Bacteriol., 172, 6704-12), fungi (Schmidhauser et al., 1990, Mol. Cell. Biol. 10, 5064-70), cyanobacteria (Chamovitz et al., 1990, Z. Naturforsch, 45c, 482-86; Cunningham et al., 1994) and higher plants (Bartley et al., Proc. Natl. Acad. Sci USA 88, 6532-36; Martinez-Ferez & Vioque, 1992, Plant Mol. Biol. 18, 981-83). Many of the isolated enzymes show a great diversity in structure, function and inhibitory properties between sources. For example, phytoene desaturases from the cyanobacterium Synechococcus and from higher plants and green algae carry out a two-step desaturation to yield ζ -carotene as a reaction product. In plants and cyanobacteria a second enzyme (ζ carotene desaturase), similar in amino acid sequence to the phytoene desaturase, catalyzes two additional desaturations to yield lycopene. In contrast, a single desaturase enzyme from Erwinia herbicola and from other bacteria introduces all four double bonds required to form lycopene. The Erwinia and other bacterial desaturases bear little amino acid sequence similarity to the plant and cyanobacterial desaturase enzymes, and are thought to be of unrelated ancestry. Therefore, even with a gene in hand from one source, it may be difficult to identify a gene encoding an enzyme of similar function in another organism. In particular, the sequence similarity between certain of the prokaryotic and eukaryotic genes encoding enzymes of carotenoid biosynthesis is quite low.

30

25

Further, the mechanism of gene expression in prokaryotes and eukaryotes appears to differ sufficiently such that one cannot expect that an isolated eukaryotic gene will be properly expressed in a prokaryotic host.

The difficulties in isolating genes encoding enzymes with similar functions is exemplified by recent efforts to isolate the gene encoding the enzyme that catalyzes the formation of β -carotene from the acyclic precursor lycopene. Although a gene encoding an enzyme with this function had been isolated from a bacterium, it had not been isolated from any photosynthetic procaryote or from any eukaryotic organism. The isolation and characterization of the enzyme catalyzing formation of β -carotene in the cyanobacterium Synechococcus PCC7942 was described by the present inventors and others (Cunningham et al., 1993 and 1994). The amino acid sequence similarity of the cyanobacterial enzyme to the various bacterial lycopene β -cyclases is so low (ca. 18-25% overall; Cunningham et al., 1994) that there is much uncertainty as to whether they share a common ancestry or, instead, represent an example of convergent evolution.

The need remains for the isolation of eukaryotic and prokaryotic genes and cDNAs encoding polypeptides involved in the carotenoid biosynthetic pathway, including those encoding a lycopene ϵ -cyclase, IPP isomerase and β -carotene hydroxylase. There remains a need for methods to enhance the production of carotenoids, to alter the composition of carotenoids, and to reduce or eliminate carotenoid production. There also remains a need in the art for methods for screening for genes and cDNAs encoding enzymes of carotenoid biosynthesis and metabolism.

SUMMARY OF THE INVENTION

20

5

10

15

Accordingly, a first object of this invention is to provide purified and/or isolated nucleic acids which encode enzymes involved in carotenoid biosynthesis; in particular, lycopene ϵ -cyclase, IPP isomerase and β -carotene hydroxylase.

A second object of this invention is to provide purified and/or isolated nucleic acids which encode enzymes which produce novel or uncommon carotenoids.

25

A third object of the present invention is to provide vectors containing said genes.

A fourth object of the present invention is to provide hosts transformed with said vectors.

Another object of the present invention is to provide hosts which accumulate novel or uncommon carotenoids or which accumulate greater amounts of specific or total carotenoids.

30

Another object of the present invention is to provide hosts with inhibited and/or altered carotenoid production.

WO 99/63055

Another object of this invention is to secure the expression of eukaryotic carotenoid-related genes in a recombinant prokaryotic host.

Yet another object of the present invention is to provide a method for screening for eukaryotic and prokaryotic genes and cDNAs which encode enzymes involved in carotenoid biosynthesis and metabolism.

An additional object of the invention is to provide a method for manipulating carotenoid biosynthesis in photosynthetic organisms by inhibiting the synthesis of certain enzymatic products to cause accumulation of precursor compounds.

Another object of the invention is to provide modified lycopene ϵ -cyclase, IPP isomerase and β -carotene hydroxylase.

These and other objects of the present invention have been realized by the present inventors as described below.

A subject of the present invention is an isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity and having the amino acid sequence of SEQ ID NOS: 2, 4, 14-21 or 23-27.

The invention also includes vectors which comprise any of the nucleic acid sequences listed above, and host cells transformed with such vectors.

Another subject of the present invention is a method of producing or enhancing the production of a carotenoid in a host cell, comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence to produce the protein.

Yet another subject of the present invention is a method of modifying the production of carotenoids in a host cell, the method comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which produces an RNA and/or encodes for a protein which modifies lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity, relative to an untransformed host cell, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence in the host cell to modify the production of the carotenoids in the host cell, relative to the untransformed host cell.

5

10

15

20

25

PCT/US99/12121

WO 99/63055

5

10

15

20

25

30

The present invention also includes a method of expressing, in a host cell, a heterologous nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity, the method comprising inserting into the host cell a vector comprising the heterologous nucleic acid sequence, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence.

Also included is a method of expressing, in a host cell, a heterologous nucleic acid sequence which encodes for a protein which modifies lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity in the host cell, relative to an untransformed host cell, the method comprising inserting into the host cell a vector comprising the heterologous nucleic acid sequence, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence.

Another subject of the present invention is a method for screening for genes and cDNAs which encode enzymes involved in carotenoid biosynthesis and metabolism.

BRIEF DESCRIPTION OF THE DRAWINGS

A more complete appreciation of the invention and many of the attendant advantages thereof will be readily obtained as the same becomes better understood by reference to the following detailed description when considered in connection with the accompanying drawings, wherein:

Figure 1 is a schematic representation of the putative pathway of β -carotene biosynthesis in cyanobacteria, algae and plants. The enzymes catalyzing various steps are indicated at the left. Target sites of the bleaching herbicides NFZ and MPTA are also indicated at the left. Abbreviations: DMAPP, dimethylallyl pyrophosphate; FPP, farnesyl pyrophosphate; GGPP, geranylgeranyl pyrophosphate; GPP, geranyl pyrophosphate; IPP, isopentenyl pyrophosphate; LCY, lycopene cyclase; MVA, mevalonic acid; MPTA, 2-(4-methylphenoxy)triethylamine hydrochloride; NFZ, norflurazon; PDS, phytoene desaturase; PSY, phytoene synthase; ZDS, ζ -carotene desaturase; PPPP, prephytoene pyrophosphate.

Figure 2 depicts possible routes of synthesis of cyclic carotenoids and common plant and algal xanthophylls (oxycarotenolds) from neurosporene. Demonstrated activities of the β - and ϵ -cyclase enzymes of A. thaliana are indicated by bold arrows labelled with β or ϵ respectively. A bar below the arrow leading to ϵ -carotene indicates that the enzymatic

activity was examined but no product was detected. The steps marked by an arrow with a dotted line have not been specifically examined. Conventional numbering of the carbon atoms is given for neurosporene and α -carotene. Inverted triangles (∇) mark positions of the double bonds introduced as a consequence of the desaturation reactions.

Figure 3 depicts the carotene endgroups which are found in plants.

Figure 4 is a DNA sequence and the predicted amino acid sequence of a lycopene ϵ -cyclase cDNA isolated from A. thaliana (SEQ ID NOS: 1 and 2). These sequences were deposited under Genbank accession number U50738. This cDNA is incorporated into the plasmid pATeps.

Figure 5 is a DNA sequence encoding the β -carotene hydroxylase isolated from A. thaliana (SEQ ID NO: 3). This cDNA is incorporated into the plasmid pATOHB.

Figure 6 is an alignment of the predicted amino acid sequences of A. thaliana β-carotene hydroxylase (SEQ ID NO: 4) with those of the bacterial β-carotene hydroxylase enzymes from Alicalgenes sp. (SEQ ID NO: 5) (Genbank D58422), Erwinia herbicola Eho10 (SEQ ID NO.: 6) (GenBank M872280), Erwinia uredovora (SEQ ID NO.: 7) (GenBank D90087) and Agrobacterium aurianticum (SEQ ID NO.: 8) (GenBank D58420). A consensus sequence is also shown. All five genes are identical where a capital letter appears in the consensus. A lowercase letter indicates that three of five, including A. thaliana, have the identical residue. TM; transmembrane.

Figure 7 is a DNA sequence of a cDNA encoding an IPP isomerase isolated from A. thaliana (SEQ ID NO: 9). This cDNA is incorporated into the plasmid pATDP5.

Figure 8 is a DNA sequence of a second cDNA encoding another IPP isomerase isolated from A. thaliana (SEQ ID NO: 10). This cDNA is incorporated into the plasmid pATDP7.

Figure 9 is a DNA sequence of a cDNA encoding an IPP isomerase isolated from *Haematococcus pluvialis* (SEQ ID NO: 11). This cDNA is incorporated into the plasmid pHP04.

Figure 10 is a DNA sequence of a second cDNA encoding another IPP isomerase isolated from *Haematococcus pluvialis* (SEQ ID NO: 12). This cDNA is incorporated into the plasmid pHP05.

Figure 11 is an alignment of the amino acid sequences predicted by IPP isomerase cDNAs isolated from A. thaliana (SEQ ID NO.: 16 and 18), H. pluvialis (SEQ ID NOS.: 14

5

10

15

20

25

and 15), Clarkia breweri (SEQ ID NO.: 17) (See, Blanc & Pichersky, Plant Physiol. (1995) 108:855; Genbank accession no. X82627) and Saccharomyces cerevisiae (SEQ ID NO.: 19) (Genbank accession no. J05090).

Figure 12 is a DNA sequence of the cDNA encoding an IPP isomerase isolated from *Tagetes erecta* (marigold; SEQ ID NO: 13). This cDNA is incorporated into the plasmid pPMDP1. xxx's denote a region not originally sequenced. Figure 21A shows the complete marigold sequence.

Figure 13 is an alignment of the consensus sequence of four plant β -cyclases (SEQ ID NO.: 20) with the A. thaliana lycopene ϵ -cyclase (SEQ ID NO.: 21). A capital letter in the plant β consensus is used where all four β -cyclase genes predict the same amino acid residue in this position. A small letter indicates that an identical residue was found in three of the four. Dashes indicate that the amino acid residue was not conserved and dots in the sequence denote a gap. A consensus for the aligned sequences is given, in capital letters below the alignment, where the β - and ϵ -cyclases have the same amino acid residue. Arrows indicate some of the conserved amino acids that will be used as junction sites for construction of chimeric cyclases with novel enzymatic activities. Several regions of interest including a sequence signature indicative of a dinucleotide-binding motif and two predicted transmembrane (TM) helical regions are indicated below the alignment and are underlined.

Figure 14 shows the nucleotide (SEQ ID NO:22) and amino acid sequences (SEQ ID NO:23) of the *Adonis palaestina* (pheasant's eye) ϵ -cyclase cDNA #5.

Figure 15A shows the nucleotide (SEQ ID NO:24) and amino acid sequences (SEQ ID NO:25) of a potato ε-cyclase cDNA. Figure 15B shows the amino acid sequence (SEQ ID NO:26) of a chimeric lettuce/potato lycopene ε-cyclase. Amino acids in lower case are from the lettuce cDNA and those in upper case are from the potato cDNA. The product of this chimeric cDNA has e-cyclase activity and converts lycopene to the monocyclic δ-carotene.

Figure 16 shows a comparison between the amino acid sequences of the *Arabidopsis* ϵ -cyclase (SEQ ID NO:27) and the potato ϵ -cyclase (SEQ ID NO:25).

Figure 17A shows the nucleotide sequence of the *Adonis palaestina* Ipi1 (SEQ ID NO:28) and Figure 17B shows the nucleotide sequence of the *Adonis palaestina* Ipi2 (SEQ ID NO: 29).

5

10

15

20

25

Figure 18A shows the nucleotide sequence of the *Haematoccus pluvialis* Ipi1 (SEQ ID NO:11) and Figure 18B shows the nucleotide sequence of the *Haematoccus pluvialis* Ipi2 (SEQ ID NO:30).

Figure 19A shows the nucleotide sequence of the Lactuca sativa (romaine lettuce)

Ipi1 (SEQ ID NO:31) and Figure 19B shows the nucleotide sequence of the Lactuca sativa

Ipi2 (SEQ ID NO: 32).

Figure 20 shows the nucleotide sequence of the *Chlamydomonas reinhardtii* Ipi1 (SEQ ID NO:33).

Figure 21A shows the nucleotide sequence of the *Tagetes erecta* (marigold) Ipi1 (SEQ ID NO:34) and Figure 21B shows the nucleotide sequence of the *Oryza sativa* (rice) Ipi1 (SEQ ID NO:35).

Figure 22 shows a amino acid sequence alignment of various plant and green algal isopentenyl isomerases (IPI) (SEQ ID NOS:16, 36-45).

Figure 23 shows a comparison between Adonis palaestina ϵ -cyclase cDNA #3 and cDNA #5 nucleotide sequences.

Figure 24 shows a comparison between *Adonis palaestina* ε-cyclase cDNA #3 and cDNA #5 predicted amino acid sequences.

Figure 25 shows a sequence alignment of various plant β - and ϵ -cyclases. Those sequences outlined in grey denote identical sequences among the ϵ -cyclases. Those sequences outlined in black denote identical sequences among both the β - and ϵ -cyclases.

Figure 26 shows a sequence alignment of the plant ϵ -cyclases from Figure 25. Those sequences outlined in black denote identical sequences among the ϵ -cyclases.

Figure 27 is a dendrogram or "tree" illustrating the degree of amino acid sequence similarity for various lycopene β - and ϵ -cyclases.

Figure 28 shows a comparison between Arabidopsis ϵ -cyclase and lettuce ϵ -cyclase predicted amino acid sequences.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention includes an isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity and having the amino acid sequence of SEQ ID NOS: 2, 4, 14-21, 23 or 25-27. Nucleic acids encoding lycopene ϵ -cyclase, β -carotene hydroxylase and IPP

5

10

15

20

25

PCT/US99/12121

ون مو

WO 99/63055

5

10

15

20

25

30

isomerases have been isolated from several genetically distant sources.

The present inventors have isolated nucleic acids encoding the enzyme IPP isomerase, which catalyzes the reversible conversion of isopentenyl pyrophosphate (IPP) to dimethylallyl pyrophosphate (DMAPP). IPP isomerase cDNAs were isolated from the plants A. thaliana, Tagetes erecta (marigold), Adonis palaestina (pheasant's eye), Lactuca sativa (romaine lettuce) and from the green algae H. pluvialis and Chlamydomonas reinhardtii. Alignments of the amino acid sequences predicted by some of these cDNAs are shown in Figures 12 and 22. Plasmids containing some of these cDNAs were deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville MD 20852 on March 4, 1996 under ATCC accession numbers 98000 (pHP05 - H. pluvialis); 98001 (pMDP1 - marigold); 98002 (pATDP7 - A. thaliana) and 98004 (pHP04 - H. pluvialis).

The present inventors have also isolated nucleic acids encoding the enzyme β -carotene hydroxylase, which is responsible for hydroxylating the β -endgroup in carotenoids. The nucleic acid of the present invention is shown in SEQ ID NO: 3 and Figure 5. The full length cDNA product hydroxylates both end groups of β -carotene as do products of cDNAs which encode proteins truncated by up to 50 amino acids from the N-terminus. Products of genes which encode proteins truncated between about 60-110 amino acids from the N-terminus preferentially hydroxylate only one ring. A plasmid containing this gene was deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville MD 20852 on March 4, 1996 under ATCC accession number 98003 (pATOHB - A. thaliana).

The present inventors have also isolated nucleic acids encoding the enzyme lycopene ϵ -cyclase, which is responsible for the formation of ϵ -endgroups in carotenoids. The A. thaliane ϵ -cyclase adds an ϵ ring to only one end of the symmetrical lycopene while the related β -cyclase adds a ring at both ends. The A. thaliana cDNA of the present invention is shown in Figure 4 and SEQ ID NO: 1. A plasmid containing this gene was deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville MD 20852 on March 4, 1996 under ATCC accession number 98005 (pATeps - A. thaliana).

In addition, lycopene ϵ -cyclases have been identified in lettuce and in *Adonis* palaestina (cDNA #5) which encode enzymes that convert lycopene to the bicyclic ϵ -carotene (ϵ , ϵ -carotene). An additional cDNA from *Adonis palaestina* (cDNA #3) encodes a lycopene ϵ -cyclase which converts lycopene into δ -carotene (ϵ , ψ -carotene) and differs from the lycopene ϵ -cyclase which forms bicyclic ϵ -carotene (ϵ , ϵ -carotene) by only 5 amino acids.

10

15

20

25

30

One or more of these amino acids may be modified by alteration of the nucleotide sequence in the #5 cDNA to obtain an enzyme which forms the bicyclic ϵ, ϵ -carotene. The sequences of the *Adonis palaestina* and *Arabidopsis thaliana* ϵ -cyclases have about 70% nucleotide identity and about 72% amino acid identity.

Initial experiments by the inventors with chimeric genes indicated that the part of the ϵ -cyclase which is responsible for adding 2ϵ rings to form ϵ, ϵ -carotene is the carboxy terminal portion of the gene. The lettuce ϵ -cyclase adds two ϵ rings to form ϵ, ϵ -carotene. A DNA encoding a partial potato ϵ -cyclase (missing its amino terminal portion), when combined with an amino terminal region from the lettuce ϵ -cyclase gene, produces a monocyclic δ -carotene (ϵ, ψ -carotene). With the discovery of the differences between the Adonis palaestina clone #3 and clone #5, the specific amino acids responsible for the addition of an extra ϵ ring have been identified (Figure 24). Specifically, amino acid 55 is Thr in clone #3 and Ser in clone #5, amino acid 210 is Asn in clone #3 and Asp in clone #5, amino acid 231 is Asp in clone #3 and Glu in clone #5, amino acid 352 is Ile in clone #3 and Val in clone #5, and amino acid 524 is Lys in clone #3 and Arg in clone #5. It can be appreciated that these changes are quite conservative, as only one change, at amino acid 210, changes the charge of the protein.

Thus, it is clear that the nucleic acids of the invention encoding the enzymes as presently disclosed may be altered to increase a particularly desirable property of the enzyme, to change a property of the enzyme, or to diminish an undesirable property of the enzyme. Such modifications can be by deletion, substitution, or insertion of one or more amino acids, and can be performed by routine enzymatic manipulation of the nucleic acid encoding the enzyme (such as by restriction enzyme digestion, removal of nucleotides by mung bean nuclease or *Bal31*, insertion of nucleotides by Klenow fragment, and by religation of the ends), by site-directed mutagenesis, or may be accidental, such as by low fidelity PCR or those obtained through mutations in hosts that are producers of the enzymes. These techniques as well as other suitable techniques are well known in the art.

Mutations can be made in the nucleic acids of the invention such that a particular codon is changed to a codon which codes for a different amino acid. Such a mutation is generally made by making the fewest nucleotide changes possible. A substitution mutation of this sort can be made to change an amino acid in the resulting protein in a non-conservative manner (i.e., by changing the codon from an amino acid belonging to a grouping

10

15

20

25

30 .

of amino acids having a particular size or characteristic to an amino acid belonging to another grouping) or in a conservative manner (i.e., by changing the codon from an amino acid belonging to a grouping of amino acids having a particular size or characteristic to an amino acid belonging to the same grouping). Such a conservative change generally leads to less change in the structure and function of the resulting protein. A non-conservative change is more likely to alter the structure, activity or function of the resulting protein. The present invention should be considered to include sequences containing conservative changes which do not significantly alter the activity or binding characteristics of the resulting protein.

The following is one example of various groupings of amino acids:

Amino acids with nonpolar R groups: Alanine, Valine, Leucine, Isoleucine, Proline, Phenylalanine, Tryptophan and Methionine.

Amino acids with uncharged polar R groups: Glycine, Serine, Threonine, Cysteine, Tyrosine, Asparagine and Glutamine.

Amino acids with charged polar R groups (negatively charged at Ph 6.0): Aspartic acid and Glutamic acid.

Basic amino acids (positively charged at pH 6.0): Lysine, Arginine and Histidine.

Another grouping may be those amino acids with phenyl groups: Phenylalanine, Tryptophan and Tyrosine.

Another grouping may be according to molecular weight (i.e., size of R groups). Particularly preferred substitutions are:

- Lys for Arg and vice versa such that a positive charge may be maintained;
- Glu for Asp and vice versa such that a negative charge may be maintained;
- Ser for Thr such that a free -OH can be maintained; and
- Gln for Asn such that a free NH₂ can be maintained.

Amino acid substitutions may also be introduced to substitute an amino acid with a particularly preferable property. For example, a Cys may be introduced to provide a potential site for disulfide bridges with another Cys. A His may be introduced as a particularly "catalytic" site (i.e., His can act as an acid or base and is the most common amino acid in biochemical catalysis). Pro may be introduced because of its particularly planar structure, which induces β -turns in the protein's structure.

It is clear that certain modifications of SEQ ID NOS: 2, 4, 14-21, 23 or 25-27 can take place without destroying the activity of the enzyme. It is noted especially that truncated

10

15

20

25

30

versions of the nucleic acids of the invention are functional. For example, several amino acids (from 1 to about 120) can be deleted from the N-terminus of the lycopene €-cyclases of the invention, and a functional protein can still be produced. This fact is made especially clear from Figure 25, which shows a sequence alignment of several plant €-cyclases. As can be seen from Figure 25, there is an enormous amount of sequence disparity between amino acid sequences 2 to about 50-70 (depending on the particular sequence, since gaps are present). There is less, but also a substantial amount of, sequence dissimilarity between about 50-70 to about 90-120 (depending on the particular sequence). Thereafter, the sequences are fairly conserved, except for small pockets of dissimilarity between about 275-295 to about 285-305 (depending on the particular sequence), and between about 395-415 to about 410-430 (depending on the particular sequence).

The present inventors have found that the amount of the 5' region present in the nucleic acids of the invention can alter the activity of the enzyme. Instead of diminishing activity, truncating the 5' region of the nucleic acids of the invention may result in an enzyme with a different specificity. Thus, the present invention relates to nucleic acids and enzymes encoded thereby which are truncated to within 0-50, preferably 0-25, codons of the 5' initiation codon of their prokaryotic counterparts as determined by alignment maps as discussed below.

For example, when the cDNA encoding A. thaliana β -carotene hydroxylase was truncated, the resulting enzyme catalyzed the formation of β -cryptoxanthin as the major product and zeaxanthin as minor product; in contrast to its normal production of zeaxanthin.

The present invention is intended to include those nucleic acid and amino acid sequences in which substitutions, deletions, additions or other modifications have taken place, as compared to SEQ ID NOS: 2, 4, 14-21, 23 or 25-27, without destroying the activity of the enzyme. Preferably, the substitutions, deletions, additions or other modifications take place at the 5' end, or any other of those positions which already show dissimilarity between any of the presently disclosed amino acid sequences (see also Figure 25) or other amino acid sequences which are known in the art and which encode the same enzyme (i.e., lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase).

In each case, nucleic acid and amino acid sequence similarity and identity is measured using sequence analysis software, for example, the Sequence Analysis, Gap, or BestFit software packages of the Genetics Computer Group (University of Wisconsin Biotechnology



Center, 1710 University Avenue, Madison, Wisconsin 53705), MEGAlign (DNAStar, Inc., 1228 S. Park St., Madison, Wisconsin 53715), or MacVector (Oxford Molecular Group, 2105 S. Bascom Avenue, Suite 200, Campbell, California 95008). Such software uses algorithms to match similar sequences by assigning degrees of identity to various substitutions, deletions, and other modifications, and includes detailed instructions as to useful parameters, etc., such that those of routine skill in the art can easily compare sequence similarities and identities. An example of a useful algorithm in this regard is the algorithm of Needleman and Wunsch, which is used in the Gap program discussed above. This program finds the alignment of two complete sequences that maximizes the number of matches and minimizes the number of gaps. Another useful algorithm is the algorithm of Smith and Waterman, which is used in the BestFit program discussed above. This program creates an optimal alignment of the best segment of similarity between two sequences. Optimal alignments are found by inserting gaps to maximize the number of matches using the local homology algorithm of Smith and Waterman.

15

10

5

Conservative (i.e. similar) substitutions typically include substitutions within the following groups: glycine and alanine; valine, isoleucine and leucine; aspartic acid, glutamic acid, asparagine and glutamine; serine and threonine; lysine and arginine; and phenylalanine and tyrosine. Substitutions may also be made on the basis of conserved hydrophobicity or hydrophilicity (see Kyte and Doolittle, *J. Mol. Biol.* 157: 105-132 (1982)), or on the basis of the ability to assume similar polypeptide secondary structure (see Chou and Fasman, *Adv. Enzymol.* 47: 45-148 (1978)).

25

20

If comparison is made between nucleotide sequences, preferably the length of comparison sequences is at least 50 nucleotides, more preferably at least 60 nucleotides, at least 75 nucleotides or at least 100 nucleotides. It is most preferred if comparison is made between the nucleic acid sequences encoding the enzyme coding regions necessary for enzyme activity. If comparison is made between amino acid sequences, preferably the length of comparison is at least 20 amino acids, more preferably at least 30 amino acids, at least 40 amino acids or at least 50 amino acids. It is most preferred if comparison is made between the amino acid sequences in the enzyme coding regions necessary for enzyme activity.

30

It should be appreciated that also within the scope of the present invention are nucleic acid sequences encoding lycopene ϵ -cyclases, IPP isomerases and β -carotene hydroxylases

10

15

20

25

30

which code for enzymes having the same amino acid sequence as SEQ ID NOS: 2, 4, 14-21, 23 or 25-27, but which are degenerate to the nucleic acids specifically disclosed herein.

The amino acid residues described herein are preferred to be in the "L" isomeric form. However, residues in the "D" isomeric form can be substituted for any L-amino acid residue, as long as the desired functional property of immunoglobulin-binding is retained by the polypeptide.

In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Sambrook et al, "Molecular Cloning: A Laboratory Manual" (1989); "Current Protocols in Molecular Biology" Volumes I-III [Ausubel, R. M., ed. (1994)]; "Cell Biology: A Laboratory Handbook" Volumes I-III [J. E. Celis, ed. (1994))]; "Current Protocols in Immunology" Volumes I-III [Coligan, J. E., ed. (1994)]; "Oligonucleotide Synthesis" (M.J. Gait ed. 1984); "Nucleic Acid Hybridization" [B.D. Hames & S.J. Higgins eds. (1985)]; "Transcription And Translation" [B.D. Hames & S.J. Higgins, eds. (1984)]; "Animal Cell Culture" [R.I. Freshney, ed. (1986)]; "Immobilized Cells And Enzymes" [IRL Press, (1986)]; B. Perbal, "A Practical Guide To Molecular Cloning" (1984).

The present invention also includes vectors. Suitable vectors according to the present invention comprise a nucleic acid of the invention encoding an enzyme involved in carotenoid biosynthesis or metabolism and a suitable promoter for the host, and can be constructed using techniques well known in the art (for example Sambrook et al., Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; Ausubel et al., Current Protocols in Molecular Biology, Greene Publishing and Wiley Interscience, New York, 1991). Suitable vectors for eukaryotic expression in plants are described in Frey et al., Plant J. (1995) 8(5):693 and Misawa et al, 1994a; incorporated herein by reference. Suitable vectors for prokaryotic expression include pACYC184, pUC119, and pBR322 (available from New England BioLabs, Bevery, MA) and pTrcHis (Invitrogen) and pET28 (Novagen) and derivatives thereof. The vectors of the present invention can additionally contain regulatory elements such as promoters, repressors, selectable markers such as antibiotic resistance genes, etc.

The nucleic acids encoding the carotenoid enzymes as described above, when cloned into a suitable expression vector, can be used to overexpress these enzymes in a plant

10

15

20

25

30

expression system or to inhibit the expression of these enzymes. For example, a vector containing the gene encoding lycopene ϵ -cyclase can be used to increase the amount of α -carotene and carotenoids derived from α -carotene (such as lutein and α -cryptoxanthin) in an organism and thereby alter the nutritional value, pharmacology and visual appearance value of the organism.

Therefore, the present invention includes a method of producing or enhancing the production of a carotenoid in a host cell, relative to an untransformed host cell, the method comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence to produce the protein.

The present invention also includes a method of modifying the production of carotenoids in a host cell, the method comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which produces an RNA and/or encodes for a protein which modifies lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity, relative to an untransformed host cell, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence in the host cell to modify the production of the carotenoids in the host cell, relative to the untransformed host cell.

The term "modifying the production" means that the amount of carotenoids produced in the host cell can be enhanced, reduced, or left the same, as compared to the untransformed host cell. In accordance with one embodiment of the present invention, the make-up of the carotenoids (i.e., the specific carotenoids produced) is changed vis a vis each other, and this change in make-up may result in either a net gain, net loss, or no net change in the total amount of carotenoids produced in the cell. In accordance with another embodiment of the present invention, the production or the biochemical activity of the carotenoids (or the enzymes which catalyze their formation) is enhanced by the insertion of an enzyme-encoding nucleic acid of the invention. In yet another embodiment of the invention, the production or the biochemical activity of the carotenoids (or the enzymes which catalyze their formation) may be reduced or inhibited by a number of different approaches available to those skilled in the art, including but not limited to such methodologies or approaches as anti-sense (e.g.,

WO 99/63055 PCT/US99/12121

Gray et al (1992) Plant Mol. Biol. 19:69-87), ribozymes (e.g., Wegener et al (1994) Mol. Gen. Genet. 245:465-470), co-suppression (e.g., Fray and Grierson (1993) Plant Mol. Biol. 22:589-602), targeted disruption of the gene (e.g., Schaefer et al. (1997) Plant J. 11:1195-1206), intracellular antibodies (e.g., Rondon and Marasco (1997) Ann. Rev. Microbiol. 51:257-283) or whatever other approaches rely on the knowledge or availability of the nucleic acid or amino acid sequences of the invention and/or portions thereof, to thereby reduce accumulation of carotenoids with ϵ rings and compounds derived from them (for ϵ -cyclase inhibition), or carotenoids with hydroxylated β rings and compounds derived from them (for β -hydroxylase inhibition), or, in the case if IPP isomerase, accumulation of any isoprenoid compound.

Preferably, at least a portion of the nucleic acid sequences used in the methods, vectors and host cells of the invention codes for an enzyme having an amino acid sequence which is at least 85% identical, preferably at least 90%, at least 95% or completely identical to SEQ ID NOS: 2, 4, 14-21, 23 or 25-27. Sequence identity is determined as noted above. Preferably, sequence additions, deletions or other modifications are made as indicated above, so as to not affect the function of the particular enzyme.

In a preferred embodiment, vectors are manufactured which contain a DNA encoding a eukaryotic IPP isomerase upstream of a DNA encoding a second eukaryotic carotenoid enzyme. The inventors have discovered that inclusion of an IPP isomerase gene increases the supply of substrate for the carotenoid pathway; thereby enhancing the production of carotenoid endproducts, as compared to a host cell which is not transformed with such a vector. This is apparent from the much deeper pigmentation in carotenoid-accumulating colonies of *E. coli* which also contain one of the aforementioned IPP isomerase genes when compared to colonies that lack this additional IPP isomerase gene. Similarly, a vector comprising an IPP isomerase gene can be used to enhance production of any secondary metabolite of dimethylallyl pyrophosphate and/or isopentenyl pyrophosphate (such as isoprenoids, steroids, carotenoids, etc.). The term "isoprenoid" is intended to mean any member of the class of naturally occurring compounds whose carbon skeletons are composed, in part or entirely, of isopentyl C₃ units. Preferably, the carbon skeleton is of an essential oil, a fragrance, a rubber, a carotenoid, or a therapeutic compound, such as paclitaxel.

A vector containing the cDNA encoding a lycopene ϵ -cyclase of the invention, preferably the lettuce lycopene ϵ -cyclase or Adonis ϵ -cyclase #5, can be used to increase the

5

10

15

20

25

30

10

15

20

25

30



amount of bicyclic ε-carotene in an organism and thereby alter the nutritional value, pharmacology and visual appearance value of the organism. In addition, the transformed organism can be used in the formulation of therapeutic agents, for example in the treatment of cancer (see Mayne et al (1996) FASEB J. 10:690-701; Tsushima et al (1995) Biol. Pharm. Bull. 18:227-233).

An antisense strand of a nucleic acid of the invention can be inserted into a vector. For example, the lycopene ϵ -cyclase gene can be inserted into a vector and incorporated into the genomic DNA of a host, thereby inhibiting the synthesis of ϵ , β -carotenoids (lutein and α -carotene) and enhancing the synthesis of β , β -carotenoids (zeaxanthin and β -carotene).

The present invention also relates to novel enzymes which are encoded by the amino acid sequences of the invention, or portions thereof.

The present invention also relates to novel enzymes which can transform known carotenoids into novel or uncommon products. Currently ϵ -carotene (see Figure 2) and γ -carotene are commonly produced only in minor amounts. As described below, an enzyme can be produced which transforms lycopene to γ -carotene and lycopene to ϵ -carotene. With these products in hand, bulk synthesis of other carotenoids derived from them are possible. For example, ϵ -carotene can be hydroxylated to form lactucaxanthin, an isomer of lutein (one ϵ and one β ring) and zeaxanthin (two β rings) where both endgroups are, instead, ϵ rings.

In addition to novel enzymes produced by truncating the 5' region of known enzymes, as discussed above, novel enzymes which can participate in the formation of unusual carotenoids can be formed by replacing portions of one gene with an analogous sequence from a structurally related gene. For example, β -cyclase and ϵ -cyclase are structurally related (see Figure 13). By replacing a portion of β -lycopene cyclase with the analogous portion of ϵ -cyclase, an enzyme which produces γ -carotene will be produced (one β endgroup). Further, by replacing a portion of the lycopene ϵ -cyclase with the analogous portion of β -cyclase, an enzyme which produces ϵ -carotene will be produced (with some exceptions, such as the lettuce ϵ -cyclase, plant ϵ -cyclases normally produce a compound with one ϵ -endgroup, δ -carotene). Similarly, β -hydroxylase could be modified to produce enzymes of novel function by creation of hybrids with ϵ -hydroxylase.

Host systems according to the present invention can comprise any organism that already produces carotenoids or which has been genetically modified to produce carotenoids.

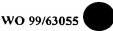
10

15

20

25

30



The IPP isomerase genes are more broadly applicable for enhancing production of any product dependent on DMAPP and/or IPP as a precursor.

Organisms which already produce carotenoids include plants, algae, some yeasts, fungi and cyanobacteria and other photosynthetic bacteria. Transformation of these hosts with vectors according to the present invention can be done using standard techniques such as those described in Misawa et al., (1990) supra; Hundle et al., (1993) supra; Hundle et al., (1991) supra; Misawa et al., (1991) supra; Sandmann et al., supra; and Schnurr et al., supra.

Transgenic organisms can be constructed which include the nucleic acid sequences of the present invention (Bird et al, 1991; Bramley et al, 1992; Misawa et al, 1994a; Misawa et al, 1994b; Cunningham et al, 1993). The incorporation of these sequences can allow the controlling of carotenoid biosynthesis, content, or composition in the host cell. These transgenic systems can be constructed to incorporate sequences which allow for the overexpression of the nucleic acids of the present invention. Transgenic systems can also be constructed containing antisense expression of the nucleic acid sequences of the present invention. Such antisense expression would result in the accumulation of the substrates of the substrates of the enzyme encoded by the sense strand.

A method for screening for eukaryotic genes which encode enzymes involved in carotenoid biosynthesis comprises transforming a prokaryotic host with a nucleic acid which may contain a eukaryotic or prokaryotic carotenoid biosynthetic gene; culturing said transformed host to obtain colonies; and screening for colonies exhibiting a different color than colonies of the untransformed host.

Suitable hosts include E. coli, cyanobacteria such as Synechococcus and Synechocystis, alga and plant cells. E. coli are preferred.

In a preferred embodiment, the above "color complementation" screening protocol can be enhanced by using mutants which are either (1) deficient in at least one carotenoid biosynthetic gene or (2) overexpress at least one carotenoid biosynthetic gene. In either case, such mutants will accumulate carotenoid precursors.

Prokaryotic and eukaryotic DNA or cDNA libraries can be screened in total for the presence of genes of carotenoid biosynthesis, metabolism and degradation. Preferred organisms to be screened include photosynthetic organisms.

10

15

20

25

30



E. coli can be transformed with these eukaryotic cDNA libraries using conventional methods such as those described in Sambrook et al, 1989 and according to protocols described by the vendors of the cloning vectors.

For example, the cDNA libraries in bacteriophage vectors such as lambdaZAP (Stratagene) or lambda ZIPLOX (Gibco BRL) can be excised en masse and used to transform *E.coli*.

Transformed *E. coli* can be cultured using conventional techniques. The culture broth preferably contains antibiotics to select and maintain plasmids. Suitable antibiotics include penicillin, ampicillin, chloramphenicol, etc. Culturing is typically conducted at 15-40°C, preferably at room temperature or slightly above (18-28°C), for 12 hours to 7 days.

Cultures are plated and the plates are screened visually for colonies with a different color than the colonies of the host $E.\ coli$ transformed with the empty plasmid cloning vector. For example, $E.\ coli$ transformed with the plasmid, pAC-BETA (described below), produce yellow colonies that accumulate β -carotene. After transformation with a cDNA library, colonies which contain a different hue than those formed by $E.\ coli/pAC$ -BETA would be expected to contain enzymes which modify the structure or accumulation of β -carotene. Similar $E.\ coli$ strains can be engineered which accumulate earlier products in carotenoid biosynthesis, such as lycopene, γ -carotene, etc.

Having generally described this invention, a further understanding can be obtained by reference to certain specific examples which are provided herein for purposes of illustration only and are not intended to be limiting unless otherwise specified.

EXAMPLE

I. <u>Isolation of β-carotene hydroxylase</u>

Plasmid Construction

An 8.6kb BgIII fragment containing the carotenoid biosynthetic genes of *Erwinia* herbicola was first cloned in the BamHI site of plasmid vector pACYC184 (chloramphenicol resistant), and then a 1.1kb BamHI fragment containing the *E. herbicola* β-carotene hydroxylase (*CrtZ*) was deleted. *E.coli* strains containing the resulting plasmid, pAC-BETA, accumulate β-carotene and form yellow colonies (Cunningham et al., 1994).

A full length cDNA encoding IPP isomerase of *Haematococcus pluvialis* (HP04) was first excised with *BamH*I and *Kpn*I from pBluescript SK-, and then ligated into the

WO 99/63055 PCT/US99/12121

corresponding sites of the pTrcHisA vector with high-level expression from the *trc* promoter (Invitrogen, Inc.). A fragment containing the IPP isomerase and *trc* promoter was subsequently excised with *EcoRV* and *KpnI*, treated with the Klenow fragment of DNA polymerase to produce blunt ends, and ligated in the Klenow-treated *HindIII* site of pAC-BETA. *E.coli* cells transformed with this new plasmid pAC-BETA-04 form orange colonies on LB plates (*vs.* yellow for those containing pAC-BETA) and cultures accumulate substantially more β-carotene (*ca.* two fold) than those that contain pAC-BETA.

Screening of an Arabidopsis cDNA Library

Several λ cDNA expression libraries of *Arabidopsis* were obtained from the *Arabidopsis* Biological Resource Center (Ohio State University, Columbus, OH) (Kieber et al., 1993). The λ cDNA libraries were excised *in vivo* using Stratagene's ExAssist SOLR system to produce a phagemid cDNA library wherein each phagemid contained also a gene conferring resistance to the antibiotic ampicillin.

E.coli strain DH10BZIP was chosen as the host cell for the screening and pigment production, although we have also used TOP10F' and XL1-Blue for this purpose. DH10B cells were transformed with plasmid pAC-BETA-04 and were plated on LB agar plates containing chloramphenicol at 50 µg/ml (from United States Biochemical Corporation). The phagemid Arabidopsis cDNA library was then introduced into DH10B cells already containing pAC-BETA-04. Transformed cells containing both pAC-BETA-04 and Arabidopsis cDNA library phagemids were selected on chloramphenicol plus ampicillin (150 μg/ml) agar plates. Maximum color development occurred after 3 to 7 days incubation at room temperature, and the rare bright yellow colonies were selected from a background of many thousands of orange colonies on each agar plate. Selected colonies were inoculated into 3 ml liquid LB medium containing ampicillin and chloramphenicol, and cultures were incubated at room temperature for 1-2 days, with shaking. Cells were then harvested by centrifugation and extracted with acetone in microfuge tubes. After centrifugation, the pigmented extract was spotted onto silica gel thin-layer chromatography (TLC) plates, and developed with a hexane:ether (1:1, by volume) mobile phases. B-carotene hydroxylaseencoding cDNAs were identified based on the appearance of a yellow pigment that comigrated with zeaxanthin on the TLC plates.

5

10

15

20

25

30



Subcloning and Sequencing

The plasmid containing the β-carotene hydroxylase cDNA was recovered and analyzed by standard procedures (Sambrook et al., 1989). The *Arabidopsis* β-carotene hydroxylase was sequenced completely on both strands on an automatic sequencer (Applied Biosystems, Model 373A, Version 2.0.1S). The cDNA insert of 0.95kb also was excised and ligated into the a pTrcHis vector. A *BgI*II restriction site within the cDNA was used to remove that portion of the cDNA that encodes the predicted polypeptide N terminal sequence region that is not also found in bacterial β-carotene hydroxylases (Figure 6). A BgIII-XhoI fragment was directionally cloned in BamHI-XhoI digested TrcHis vectors.

Pigment Analysis

5

10

15

20

25

30

A single colony was used to inoculate 50 ml of LB containing ampicillin and chloramphenicol in a 250-ml flask. Cultures were incubated at 28°C for 36 hours with gentle shaking, and then harvested at 5000 rpm in an SS-34 rotor. The cells were washed once with distilled H₂O and resuspended with 0.5 ml of water. The extraction procedures and HPLC were essentially as described previously (Cunningham et al, 1994).

II. Isolation and biochemical analysis of an Arabidopsis lycopene ϵ -cyclase Plasmid Construction

Construction of plasmids pAC-LYC, pAC-NEUR, and pAC-ZETA is described in Cunningham et al., (1994). In brief, the appropriate carotenoid biosynthetic genes from *Erwinia herbicola*, *Rhodobacter capsulatus*, and *Synechococcus* sp. strain PCC7942 were cloned in the plasmid vector pACYC184 (New England BioLabs, Beverly, MA). Cultures of *E. coli* containing the plasmids pAC-ZETA, pAC-NEUR, and pAC-LYC, accumulate ζ-carotene, neurosporene, and lycopene, respectively. The plasmid pAC-ZETA was constructed as follows: an 8.6-kb BglII fragment containing the carotenoid biosynthetic genes of *E. herbicola* (GenBank M87280; Hundle et al., 1991) was obtained after partial digestion of plasmid pPL376 (Perry et al., 1986; Tuveson et al., 1986) and cloned in the BamHI site of pACYC184 to give the plasmid pAC-EHER. Deletion of adjacent 0.8- and 1.1-kb BamHI-BamHI fragments (deletion Z in Cunningham et al., 1994), and of a 1.1 kB SalI-SalI fragment (deletion X) served to remove most of the coding regions for the *E. herbicola* β-carotene hydroxylase (crtZ gene) and zeaxanthin glucosyltransferase (crtX gene), respectively. The

resulting plasmid, pAC-BETA, retains functional genes for geranylgeranyl pyrophosphate synthase (crtE), phytoene synthase (crtB), phytoene desaturase (crtI), and lycopene cyclase (crtY). Cells of *E. coli* containing this plasmid form yellow colonies and accumulate β -carotene. A plasmid containing both the lycopene ϵ - and β -cyclase cDNAs of *A. thaliana* was constructed by excising the ϵ -cyclase in clone y2 as a PvuI-PvuII fragment and ligating this piece in the SnaBI site of a plasmid (pSPORT 1 from GIBCO-BRL) that already contained the β -cyclase (Cunningham et al., 1996).

Organisms and Growth Conditions

E. coli strains TOP10 and TOP10 F' (obtained from Invitrogen Corporation, San Diego, CA) and XL1-Blue (Stratagene) were grown in Luria-Bertani (LB) medium (Sambrook et al., 1989) at 37°C in darkness on a platform shaker at 225 cycles per min. Media components were from Difco (yeast extract and tryptone) or Sigma (NaCl). Ampicillin at 150 μg/mL and/or chloramphenicol at 50 μg/mL (both from United States Biochemical Corporation) were used, as appropriate, for selection and maintenance of plasmids.

Mass Excision and Color Complementation Screening of an A. thaliana cDNA Library

A size-fractionated 1-2 kB cDNA library of A. thaliana in lambda ZAPII (Kieber et al., 1993) was obtained from the Arabidopsis Biological Resource Center at The Ohio State University (stock number CD4-14). Other size fractionated libraries were also obtained (stock numbers CD4-13, CD4-15, and CD4-16). An aliquot of each library was treated to cause a mass excision of the cDNAs and thereby produce a phagemid library according to the instructions provided by the supplier of the cloning vector (Stratagene; E. coli strain XL1-Blue and the helper phage R408 were used). The titre of the excised phagemid was determined and the library was introduced into a lycopene-accumulating strain of E. coli TOP10 F' (this strain contained the plasmid pAC-LYC) by incubation of the phagemid with the E. coli cells for 15 min at 37°C. Cells had been grown overnight at 30°C in LB medium supplemented with 2% (w/v) maltose and 10 mM MgSO₄ (final concentration), and harvested in 1.5 ml microfuge tubes at a setting of 3 on an Eppendorf microfuge (5415C) for 10 min. The pellets were resuspended in 10 mM MgSO₄ to a volume equal to one-half that of the

5

10

15

20

25



initial culture volume. Transformants were spread on large (150 mm diameter) LB agar petri plates containing antibiotics to provide for selection of cDNA clones (ampicillin) and maintenance of pAC-LYC (chloramphenicol). Approximately 10,000 colony forming units were spread on each plate. Petri plates were incubated at 37·C for 16 hr and then at room temperature for 2 to 7 days to allow maximum color development. Plates were screened visually with the aid of an illuminated 3x magnifier and a low power stage-dissecting microscope for the rare, pale pinkish-yellow to deep-yellow colonies that could be observed in the background of pink colonies. A colony color of yellow or pinkish-yellow was taken as presumptive evidence of a cyclization activity. These yellow colonies were collected with sterile toothpicks and used to inoculate 3ml of LB medium in culture tubes with overnight growth at 37°C and shaking at 225 cycles/min. Cultures were split into two aliquots in microfuge tubes and harvested by centrifugation at a setting of 5 in an Eppendorf 5415C microfuge. After discarding the liquid, one pellet was frozen for later purification of plasmid DNA. To the second pellet was added 1.5 ml EtOH, and the pellet was resuspended by vortex mixing, and extraction was allowed to proceed in the dark for 15-30 min with occasional remixing. Insoluble materials were pelleted by centrifugation at maximum speed for 10 min in a microfuge. Absorption spectra of the supernatant fluids were recorded from 350-550 nm with a Perkin Elmer lambda six spectrophotometer.

Analysis of isolated clones

20

5

10

15

Eight of the yellow colonies contained β -carotene indicating that a single gene product catalyzes both cyclizations required to form the two β endgroups of the symmetrical β -carotene from the symmetrical precursor lycopene. One of the yellow colonies contained a pigment with the spectrum characteristic of δ -carotene, a monocyclic carotenoid with a single ϵ endgroup. Unlike the β cyclase, this ϵ -cyclase appears unable to carry out a second cyclization at the other end of the molecule.

25

The observation that ϵ -cyclase is unable to form two cyclic ϵ -endgroups (e.g. the bicyclic ϵ -carotene) illuminates the mechanism by which plants can coordinate and control the flow of substrate into carotenoids derived from β -carotene versus those derived from α -carotene and also can prevent the formation of carotenoids with two ϵ endgroups.

30

The availability of the A. thaliana gene encoding the ϵ -cyclase enables the directed manipulation of plant and algal species for modification of carotenoid content and

composition. Through inactivation of the ϵ -cyclase, whether at the gene level by deletion of the gene or by insertional inactivation or by reduction of the amount of enzyme formed (by such as antisense technology), one may increase the formation of β -carotene and other pigments derived from it. Since vitamin A is derived only from carotenoids with β endgroups, an enhancement of the production of β -carotene versus α -carotene may enhance nutritional value of crop plants. Reduction of carotenoids with ϵ -endgroups may also be of value in modifying the color properties of crop plants and specific tissues of these plants. Alternatively, where production of α -carotene, or pigments such as lutein that are derived from α -carotene, is desirable, whether for the color properties, nutritional value or other reason, one may overexpress the ϵ -cyclase or express it in specific tissues. Wherever agronomic value of a crop is related to pigmentation provided by carotenoid pigments the directed manipulation of expression of the ϵ -cyclase gene and/or production of the enzyme may be of commercial value.

The predicted amino acid sequence of the A. thaliana ϵ -cyclase enzyme was determined. A comparison of the amino acid sequences of the β - and ϵ -cyclase enzymes of Arabidopsis thaliana (Fig. 13) as predicted by the DNA sequence of the respective cDNAs (Fig. 4 for the ϵ -cyclase cDNA sequence), indicates that these two enzymes have many regions of sequence similarity, but they are only about 37% identical overall at the amino acid level. The degree of sequence identity at the DNA base level, only about 50%, is sufficiently low such that we and others have been unable to detect this gene by hybridization using the β cyclase as a probe in DNA gel blot experiments.

REFERENCES

Each reference cited in this application and/or listed below is hereby incorporated by reference.

25 Bird et al, 1991 Biotechnology 9, 635-639.

Bishop et al., (1995) FEBS Lett. 367, 158-162.

Bramley, P.M. (1985) Adv. Lipid Res. 21, 243-279.

Bramley, P.M. (1992) Plant J. 2, 343-349.

Britton, G. (1988). Biosynthesis of carotenoids. In Plant Pigments, T.W. Goodwin, ed. (London: Academic Press), pp. 133-182.

30

5

10

15

20

10

15

25



Britton, G. (1979) Z. Naturforsch. Section C Biosci. 34, 979-985.

Britton, G. (1995) UV/Visible spectroscopy. In Carotenoids, Vol. IB: Spectroscopy,

G. Britton, S. Liaaen-Jensen, H.P. Pfander, eds. (Basel: Birkhauser Verlag), pp. 13-62.

Bouvier et al., (1994) Plant J. 6, 45-54.

Cunningham et al., (1985) Photochem. Photobiol. 42: 295-307.

Cunningham et al., (1993) FEBS Lett. 328, 130-138.

Cunningham et al., (1994) Plant Cell 6, 1107-1121.

Cunningham et al., (1996) Plant Cell 8, 1613-1626.

Davies, B.H. (1976). Carotenoids. In Chemistry and Biochemistry of Plant Pigments,

Vol. 2, T.W. Goodwin, ed (New York: Academic Press), pp. 38-165.

Del Sal et al., (1988). Nucl. Acids Res. 16, 9878.

Demmig-Adams & Adams, (1992) Ann. Rev. Plant Physiol. Mol. Biol. 43, 599-626.

Enzell & Back, (1995) Mass spectrometry. In Carotenoids, Vol. IB: Spectroscopy, G.

Britton, S. Liaaen-Jensen, H.P. Pfander, eds. (Basel: Birkhauser Verlag), pp. 261-320.

Frank & Cogdell (1993) Photochemistry and function of carotenoids in photosynthesis. In Carotenoids in Photosynthesis. A. Young and G. Britton, eds. (London: Chapman and Hall). pp. 253-326.

Goodwin, T.W. (1980). The Biochemistry of the Carotenoids. 2nd ed, Vol. 1 (London: Chapman and Hall.

20 Horvath et al., (1972) Phytochem. 11, 183-187.

Hugueney et al., (1995) Plant J. 8, 417-424.

Hundle et al., (1991) Photochem. Photobiol. 54, 89-93.

Jensen & Jensen, (1971) Methods Enzymol. 23, 586-602.

Kargl & Quackenbush, (1960) Archives Biochem. Biophys. 88, 59-63.

Kargl et al., (1960) Proc. Am. Hort. Soc. 75, 574-578.

Kieber et al., (1993) Cell 72, 427-441.

Koyama, Y. (1991) J. Photochem. Photobiol., B, 9, 265-80.

Krinsky, N.I. (1987) Medical uses of carotenoids. In Carotenoids, N.I. Krinsky, M.M. Mathews-Roth, and R.F. Taylor, eds. (New York: Plenum), pp. 195-206.

30 Kyte & Doolittle, (1982) J. Mol. Biol. 157, 105-132.

LaRossa & Schloss, (1984) J. Biol. Chem. 259, 8753-8757.

Misawa et al., (1994a) Plant J. 6, 481-489.

÷ 1

Misawa et al., (1994b) J. Biochem, Tokyo, 116, 980-985.

Norris et al., (1995) Plant Cell 7, 2139-2149.

Pecker et al., (1996) Submitted to Plant Mol. Biol.

Perry et al., (1986) J. Bacteriol. 168, 607-612.

Persson & Argos, (1994) J. Mol. Biol. 237, 182-192.

Plumley & Schmidt, (1987) Proc. Nat. Acad. Sci. USA 83, 146-150.

Plumley & Schmidt, (1995) Plant Cell 7, 689-704.

Rossmann et al., (1974) Nature 250, 194-199.

Rock & Zeevaart (1991) Proc. Nat. Acad. Sci. USA 88, 7496-7499.

Rost et al., (1995) Protein Science 4, 521-533.

Sambrook et al., (1989) Molecular Cloning: A Laboratory Manual, 2nd edition (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Sancar, A. (1994) Biochemistry 33, 2-9.

Sander & Schneider, (1991) Proteins 9, 56-68.

Sandmann, G. (1994) Eur. J. Biochem. 223, 7-24.

Scolnik & Bartley, (1995) Plant Physiol. 108, 1342.

Siefermann-Harms, D. (1987) Physiol. Plant. 69, 561-568.

Spurgeon & Porter, (1980). Biosynthesis of carotenoids. In Biochemistry of Isoprenoid Compounds, J.W. Porter, and S.L. Spurgeon, eds. (New York: Wiley), pp. 1-122.

Tomes, M.L. (1963) Bot. Gaz. 124, 180-185.

Tomes, M.L. (1967) Genetics 56, 227-232.

Tuveson et al., (1986) J. Bacteriol. 170, 4675-4680.

Van Beeumen et al., (1991) J. Biol. Chem. 266, 12921-12931.

Weedon & Moss, (1995) Structure and Nomenclature. In Carotenoids, Vol. IB:

25 Spectroscopy, G. Britton, S. Liaaen-Jensen, H.P. Pfander, eds. (Basel: Birkhauser Verlag), pp. 27-70.

Wierenga et al., (1986) J. Mol. Biol. 187, 101-107.

Zechmeister, L. (1962) Cis-Trans Isomeric Carotenoids, Vitamins A and Arylpolyenes. Springer-Verlag, Vienna.

Having now fully described the invention, it will be apparent to one of ordinary skill in the art that many changes and modifications can be made thereto without departing from the spirit or scope of the invention as set forth herein.

5

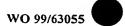
15

20

30

We claim:

- 1. An isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase enzyme activity and has an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23 or 25-27.
- 5 2. The nucleic acid sequence of claim 1, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23 or 25-27.
 - 3. A vector comprising the nucleic acid sequence of claim 1, wherein the nucleic acid sequence is operably linked to a promoter.
 - 4. A host cell which contains the vector of claim 3.
- 5. The host cell of claim 4, wherein the host cell is selected from the group consisting of a bacterial cell, an algal cell, a yeast cell and a plant cell.
 - 6. The host cell of claim 4, wherein the host cell is a photosynthetic cell.
- An isolated and/or purified protein having lycopene ε-cyclase enzyme activity and having an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23 or 25-27.
 - 8. The protein of claim 7, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23 or 25-27.





[received by the International Bureau on 15 November 1999 (15.11.99); original claims 1,2,7 and 8 amended; remaining claims unchanged (1 page)]

- 1. An isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase enzyme activity and has an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23, 25 or 26.
- 5 2. The nucleic acid sequence of claim 1, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23, 25 or 26.
 - 3. A vector comprising the nucleic acid sequence of claim 1, wherein the nucleic acid sequence is operably linked to a promoter.
 - 4. A host cell which contains the vector of claim 3.
- The host cell of claim 4, wherein the host cell is selected from the group consisting of a bacterial cell, an algal cell, a yeast cell and a plant cell.
 - 6. The host cell of claim 4, wherein the host cell is a photosynthetic cell.
- 7. An isolated and/or purified protein having lycopene ε-cyclase enzyme activity and having an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23, 25 or 26.
 - 8. The protein of claim 7, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23, 25 or 26.

FIG. 1

3
$$3 + \frac{2}{4}$$

$$4 \text{ endgroup}$$

$$8 \text{ cyclase}$$

$$3 + \frac{6}{4}$$

$$8 \text{ endgroup}$$

$$6 \text{ endgroup}$$

#¹2

4/45

FIG.4A

FIG. 4B

FIG. 4

FIG. IIA

FIG.IIB

F1G. 11

FIG. 13A

FIG. 13B

FIG. 13

FIG.14A

FIG. 14B

FIG. 14

FIG. 22A

FIG. 22 B

FIG. 22

FIG. 4A

Arabidopsis thaliana epsilon cyclase:

| | acaaaaggaaataattag attcctctttctgcttgctataccttgaca | 48 |
|-----|---|-----|
| | gaacaacataacaatggtgtaagtcttctc gctgtattcgaaattatttggaggaggaac | 108 |
| 1 | atggagtgtgttggggctaggaatttcgca gcaatggcggtttcaacatttccgtcatgg M E C V G A R N F A A M A V S T F P S W | 168 |
| 21 | agttgtcgaaggaaatttccagtggctaag agatacagctataggaatattcgcttcggt S C R R K F P V V K R Y S Y R N I R F G | 228 |
| 41 | ttgtgtagtgtcagagctagcggcggcgga agttccggtagtgagagttgtgtagcggtg L C S V R A S G G G .S S G S K S C V A V | 288 |
| 61 | agagaagatttcgctgacgaagaagatttt gcgaaagctggcggttctgagattctattt R S D F A D E E D F V E A G G S R I L F | 348 |
| 81 | gttcaaatgcagcagaacaaagatatggat gaacagtctaagcttgttgataagttgcct V Q M Q Q M K D M D S Q S K L V D K L P | 408 |
| 01 | cctatatcaactggtgatggtgctttggat catgtggttactggctgtggtcctgctggt P I S I G D G A L D K V V I G C G P A G | 468 |
| 21 | ttagccttggctgcagaatcagctaagctt ggattaaaagttggactcattggtccagat L A L A A K S A K L G L K V G L I G P D | 528 |
| 141 | cttccttttactaacaattacggtgtttgg gaagatgaattcaatgatcttgggctgcaa L P F T M M Y G V M K D K F N D L G L G | 588 |
| 161 | aaatgtattgagcatgtttggagagagact attgcgcacctggatgatgacaagcctatt K C I K K V W R S T I V Y L D D K P I | 648 |
| 181 | accattggccgtgcttatggaagagttagt cgacgtttgctccatgaggagcttttgagg T I G R A Y G R V S R R L L X E E L L R | 708 |
| 201 | aggtgtgtcgagtcaggtgtctcgtacctt agctcgaaagttgacagcataacagaagct R C V K S G V S Y L S S K V D S I T E A | 768 |
| 221 | tgtgatggccttagacttgttgcttgtgac gacaataacgtcattccctgcaggcttgcc S D G L X L V A C D D M M V I P C X L A | 828 |
| 241 | actgttgcttctggagcagcttcgggaaag ctcttgcaatacgaagttggtggacctaga T V A S G A A S G K L L Q Y X V G G P R | 888 |
| | atctatacacaactacatacaacaaaataatcaatataatca | 948 |

FIG. 4B

| 201 | V (| ٠ | ٧ | Ų | 1 | А | Y | G | ٧ | X | V | X | V | X | N | 5 | Р | Y | D | Р | |
|-----|-----------|----------|----------|-----------------|----------|-----------|-----------|-----------------|-----------|-----------|----------|-----------|-----------|-----------------|-----------------|----------|----------|------------------|-----------|----------|------|
| 281 | gato D | caa Q | atg M | gtt V | ttc P | ato M | gat D | tac Y | caga R | agat D | tat Y | act T | aac M | gag X | aaa X | gtt V | .cgg | jago S | tta L | gaa X | 1008 |
| 301 | gcto A | gag K | tat Y | .cca P | acg T | rttt F | cto L | rtac Y | gcc A | catg M | cct P | atg M | jaca T | aag K | tca S | aga R | ctc | ttc F | ttc(| gag K | 1068 |
| 321 | gaga K | aca T | tgt C | ttg L | gcc A | tca S | aaa K | igat D | gto V | atg M | ccc P | ttt F | gat D | ttg L | cta L | aaa K | acg T | aag K | ctca L | atg M | 1128 |
| 341 | ttaa I | aga P | tta V | gac G | aca G | cto S | gga L | att P | cga N | att T | cta X | iaag Q | act K | tac N | gaa L | gag A | gag F | tgg G | tcc1 A | tat A | 1188 |
| 361 | atco I | cca P | gtt V | ggt G | ggt G | tcc S | ttg L | KCZ P | aaac M | cacc T | gaa X | icaa Q | aag K | aat N | ctc L | gcc A | ttt F | ggt G | gcto A | gcc A | 1248 |
| 381 | gcta A | agc S | atg M | gta V | cat M | .ccc | :gca A | aca T | iggc G | tat Y | tca S | gtt V | gtg V | aga R | tct S | ttg L | tct S | gaa X | gcto A | cca P | 1308 |
| 401 | aaac K | at Y | gca A | tca \$ | gtc V | atc I | :gca A | gag K | jata I | icta L | aga R | gaa E | gag E | act T | acc T | aaa K | cag Q | att I | aaca N | agt S | 1368 |
| 421 | aata M | itt I | tca S | aga R | caa Q | gct A | tag W | gat D | act T | tta L | tgg W | icca P | cca P | gaa E | agg R | aaa X | aga R | cag. Q | agaç R | gca A | 1428 |
| 441 | ttc1 F | tt F | ctc L | ttt F | ggt G | ctt L | .gca A | ctc L | aga I | gtt V | caa Q | ttc F | gat D | acc T | gaa X | ggc G | att I | aga R | agct S | tc F | 1488 |
| 461 | ttco F | gt R | act T | ttc P | ttc F | cgc R | ctt L | CCa P | aaa K | itgg W | atg M | rtgg W | caa Q | 999 G | ttt F | cta L | gga G | tca S | acat T | ta L | 1548 |
| 481 | acat T | ca S | gga G | gat D | ctc L | gtt V | ctc | ttt F | gct A | tta L | tac Y | atg M | tto P | gtc V | att I | tca S | cca P | aac M | aatt M | tg L | 1608 |
| 501 | agaa R | aaa K | ggt G | ctc L | att I | aat N | .cat W | ctc L | ato I | tct S | gat D | cca P | aco T | gga G | gca A | acc T | atg M | ata I | aaaa K | acc T | 1668 |
| 521 | tato Y | | aaa K | | tga | ttt | act | tac | caa | ctc | tta | ggt | ttg | tgt | ata | tat | atg | ccg | attt | at | 1728 |
| J. | · | | | - | . | | | | | . | _1.1 | | | | | | | | | | 1700 |
| | ctg | | | | | | | | | | | | | | | | | | _ | | 1788 |
| | agaa | atc | taa | gga | gtg | atc | gaa | atg | gag | acg | gaa | acg | aaa | aga | aaa | aaa | tca | gtc | tttg | rtt | 1848 |
| | ccgt | .gg | cta | gtg | | | | | | | | | | | | | | | | | 1868 |

FIG. 5

gctctttctc ctcctcctct accgatttcc gactccgcct cccgaaatcc 51 ttatccggat tctctccgtc tcttcgattt aaacgctttt ctgtctgtta cgtcgtcgaa gaacggagac agaattctcc gattgagaac gatgagagac 101 cggagagcac gagctccaca aacgctatag acgctgagta tctggcgttg 151 201 cgtttggcgg agaaattgga gaggaagaaa tcggagaggt ccacttatct 251 aatcgctgct atgttgtcga gctttggtat cacttctatg gctgttatgg 301 ctgtttacta cagattctct tggcaaatgg agggaggtga gatctcaatg 351 ttggaaatgt ttggtacatt tgctctctct gttggtgctg ctgttggtat ggaattctgg gcaagatggg ctcatagagc tctgtggcac gcttctctat 401 451 ggaatatgca tgagtcacat cacaaaccaa gagaaggacc gtttgagcta 501 aacgatgttt ttgctatagt gaacgctggt ccagcgattg gtctcctctc 551 ttatggattc ttcaataaag gactcgttcc tggtctctgc tttggcgccg ggttaggcat aacggtgttt ggaatcgcct acatgtttgt ccacgatggt 601 651 ctcgtgcaca agcgtttccc tgtaggtccc atcgccgacg tcccttacct 701 ccgaaaggtc gccgccgctc accagctaca tcacacagac aagttcaatg 751 gtgtaccata tggactgttt cttggaccca aggaattgga agaagttgga 801 ggaaatgaag agttagataa ggagattagt cggagaatca aatcatacaa 851 aaaggcctcg ggctccgggt cgagttcgag ttcttgactt taaacaagtt 901 ttaaatccca aattctttt ttgtcttctg tcattatgat catcttaaga 951 cggtct

±2

8745

FIG. 7

ccacgggtcc gcctccccgt ttttttccga tccgatctcc ggtgccgagg 1 actcagctgt ttgttcgcgc tttctcagcc gtcaccatga ccgattctaa 51 cgatgctgga atggatgctg ttcagagacg actcatgttt gaagacgaat 101 151 gcattctcgt tgatgaaaat aatcgtgtgg tgggacatga cactaagtat 201 aactgtcatc tgatggaaaa gattgaagct gagaatttac ttcacagagc tttcagtgtg tttttattca actccaagta tgagttgctt ctccagcaac 251 ggtcaaaaac aaaggttact ttcccacttg tgtggacaaa cacttgttgc 301 agccatecte tttacegtga atecgagett attgaagaga atgtgettgg 351 tgtaagaaat gccgcacaaa ggaagctttt cgatgagctc ggtattgtag 401 451 cagaagatgt accagtegat gagtteacte cettgggaeg catgetttae 501 aaggcacctt ctgatgggaa atggggagag cacgaagttg actatctact 551 cttcatcgtg cgggatgtga agcttcaacc aaacccagat gaagtggctg 601 agatcaagta cgtgagcagg gaagagctta aggagctggt gaagaaagca 651 gatgetggeg atgaagetgt gaaactatet ceatggttea gattggtggt 701 ggataatttc ttgatgaagt ggtgggatca tgttgagaaa ggaactatca 751 ctgaagctgc agacatgaaa accattcaca agctctgaac tttccataag 801 ttttggatct tccccttccc ataataaaat taagagatga gacttttatt 851 gattacagac aaaactggca acaaaatcta ttcctaggat ttttttttgc tttttattta cttttgattc atctctagtt tagttttcat cttaaaaaaa 901 951 aaaa

FIG. 8

1 caccaatgte tgtttcttct ttatttaate teecattgat tegeeteaga 51 tototogoto titogiotto tititottot titocgattig cocatogico TCTGTCATCG ATTTCACCGA GAAAGTTACC GAATTTTCGT GCTTTCTCTG 101 GTACCGCTAT GACAGATACT AAAGATGCTG GTATGGATGC TGTTCAGAGA 151 201 CGTCTCATGT TTGAGGATGA ATGCATTCTT GTTGATGAAA CTGATCGTGT 251 TGTGGGGCAT GTCAGCAAGT ATAATTGTCA TCTGATGGAA AATATTGAAG 301 CCAAGAATTT GCTGCACAGG GCTTTTAGTG TATTTTTATT CAACTCGAAG 351 TATGAGTTGC TTCTCCAGCA AAGGTCAAAC ACAAAGGTTA CGTTCCCTCT AGTGTGGACT AACACTTGTT GCAGCCATCC TCTTTACCGT GAATCAGAGC 401 451 TTATCCAGGA CAATGCACTA GGTGTGAGGA ATGCTGCACA AAGAAAGCTT 501 CTCGATGAGC TTGGTATTGT AGCTGAAGAT GTACCAGTCG ATGAGTTCAC 551 TCCCTTGGGA CGTATGCTGT ACAAGGCTCC TTCTGATGGC AAATGGGGAG 601 AGCATGAACT TGATTACTTG CTCTTCATCG TGCGAGACGT GAAGGTTCAA CCAAACCCAG ATGAAGTAGC TGAGATCAAG TATGTGAGCC GGGAAGAGCT 651 701 GAAGGAGCTG GTGAAGAAAG CAGATGCAGG TGAGGAAGGT TTGAAACTGT 751 CACCATGGTT CAGATTGGTG GTGGACAATT TCTTGATGAA GTGGTGGGAT CATGTTGAGA AAGGAACTTT GGTTGAAGCT ATAGACATGA AAACCATCCA 801 851 CAAACTCTGA ACATCTTTTT TTAAAGTTTT TAAATCAATC AACTTTCTCT TCATCATTTT TATCTTTTCG ATGATAATAA TTTGGGATAT GTGAGACACT 901 TACAAAACTT CCAAGCACCT CAGGCAATAA TAAAGTTTGC GGCCGC

FIG. 9

| 1 | CICGGTAGCT | GGCCACAATC | GCTATTTGGA | ACCTGGCCCG | GCGGCAGTCC |
|------|------------|------------|------------|------------|------------|
| 51 | GATGCCGCGA | TGCTTCGTTC | GTTGCTCAGA | GGCCTCACGC | ATATCCCCC |
| 101 | CGTGAACTCC | GCCCAGCAGC | CCAGCTGTGC | ACACGCGCGA | CTCCAGTTTA |
| 151 | AGCTCAGGAG | CATGCAGATG | ACGCTCATGC | AGCCCAGCAT | CTCAGCCAAT |
| 201 | CTGTCGCGCG | CCGAGGACCG | CACAGACCAC | ATGAGGGGTG | CAAGCACCTG |
| 251 | GGCAGGCGGG | CAGTCGCAGG | ATGAGCTGAT | GCTGAAGGAC | GAGTGCATCT |
| 301 | TGGTGGATGT | TGAGGACAAC | ATCACAGGCC | ATGCCAGCAA | GCTGGXGTGT |
| 351 | CACAAGTTCC | TACCACATCA | GCCTGCAGGC | CTGCTGCACC | GGGCCTTCTC |
| 401 | TGTGTTCCTG | TTTGACGATC | AGGGGGGACT | GCTGCTGCAA | CAGCGTGCAC |
| 451 | GCTCAAAAAT | CACCTTCCCA | AGTGTGTGGA | CGAACACCTG | CTGCAGCCAC |
| 501 | CCTTTACATG | GGCAGACCCC | AGATGAGGTG | GACCAACTAA | GCCAGGTGGC |
| 551 | CGACGGAACA | GTACCTGGCG | CAAAGGCTGC | TGCCATCCGC | AAGTTGGAGC |
| 601 | ACGAGCTGGG | GATACCAGCG | CACCAGCTGC | CGGCAAGCGC | GTTTCGCTTC |
| 651 | CTCACGCGTT | TGCACTACTG | TGCCGCGGAC | GTGCAGCCAG | CTGCGACACA |
| 701 | ATCAGCGCTC | TGGGGCGAGC | ACGAAATGGA | CTACATCTTG | TTCATCCGG |
| 751 | CCAACGTCAC | CTTGGCGCCC | AACCCTGACG | AGGTGGACGA | AGTCAGGTAC |
| 801 | GTGACGCAAG | AGGAGCTGCG | GCAGATGATG | CAGCCGGACA | ACGGGCTGCA |
| 851 | ATGGTCGCCG | TGGTTTCGCA | TCATCGCCGC | GCGCTTCCTT | GAGCGTTGGT |
| 901 | GGGCTGACCT | GGACGCGGCC | CTAAACACTG | ACAAACACGA | GGATTGGGGA |
| 951 | ACGGTGCATC | ACATCAACGA | AGCGTGÄAAG | CAGAAGCTGC | AGGATGTGAA |
| 1001 | GACACGTCAT | GGGGTGGAAŢ | TGCGTACTTG | GCAGCTTCGT | ATCTCCTTT |
| 1051 | TCTGAGACTG | AACCTGCAGT | CAGGTCCCAC | AAGGTCAGGT | AAAATGGCTC |
| 1101 | GATAAAATGT | ACCGTCACTT | TTTGTCGCGT | ATACTGAACT | CCAAGAGGT |
| 1151 | ****** | **** | | | |

FIG. 10

| 1 | CTCGGTAGCT | GGCCACAATC | GCTATTTGGA | ACCTGGCCCG | GCGGCAGTCG |
|------------------|------------|------------|------------|------------|------------|
| 51 | GATGCCGCGA | TGCTTCGTTC | GTTGCTCAGA | GGCCTCACGC | ATATCCCGCG |
| 101 | CGTGAACTCC | GCCCAGCAGC | CCAGCTGTGC | ACACGCGCGA | CTCCAGTTTA |
| 151 | AGCTCAGGAG | CATGCAGCTG | CTTTCCGAGG | ACCGCACAGA | CCACATGAGG |
| 201 | GGTGCAAGCA | CCTGGGCAGG | CGGGCAGTCG | CAGGATGAGC | TGATGCTGAA |
| 251 | GGACGAGTGC | ATCTTGGTAG | ATGTTGAGGA | CAACATCACA | GGCCATGCCA |
| 301 | GCAAGCTGGA | GTGTCACAAG | TTCCTACCAC | ATCAGCCTGC | AGGCCTGCTG |
| 351 | CACCGGGCCT | TCTCTGTGTT | CCTGTTTGAC | GATCAGGGGC | GACTGCTGCT |
| 401 | GCAACAGCGT | GCACGCTCAA | AAATCACCTT | CCCAAGTGTG | TGGACGAACA |
| 451 | CCTGCTGCAG | CCACCCTTTA | CATGGGCAGA | CCCCAGATGA | GGTGGACCAA |
| 501 | CTAAGCCAGG | TGGCCGACGG | AACAGTACCT | GGCGCAAAGG | CTGCTGCCAT |
| 551 | CCGCAAGTTG | GAGCACGAGC | TGGGGATACC | AGCGCACCAG | CTGCCGGCAA |
| 601 | GCGCGTTTCG | CTTCCTCACG | CGTTTGCACT | ACTGTGCCGC | GGACGTGCAG |
| 651 | CCAGCTGCGA | CACAATCAGC | GCTCTGGGGC | GAGCACGAAA | TGGACTACAT |
| 701 | CTTGTTCATC | CGGGCCAACG | TCACCTTGGC | GCCCAACCCT | GACGAGGTGG |
| 751 [.] | ACGAAGTCAG | GTACGTGACG | CAAGAGGAGC | TGCGGCAGAT | GATGCAGCCG |
| 801 | GACAACGGGC | TTCAATGGTC | GCCGTGGTTT | CGCATCATCG | CCGCGCGCTT |
| 851 | CCTTGAGCGT | TGGTGGGCTG | ACCTGGACGC | GGCCCTAAAC | ACTGACAAAC |
| 901 | ACGAGGATTG | GGGAACGGTG | CATCACATCA | ACGAAGCGTG | AAGGCAGAAG |
| 951 | CTGCAGGATG | TGAAGACACG | TCATGGGGTG | GAATTGCGTA | CTTGGCAGCT |
| .001 | TCGTATCTCC | TTTTTCTGAG | ACTGAACCTG | CAGAGCTAGA | GTCAATGGTG |
| .051 | CATCATATTC | ATCGTCTCTC | TTTTGTTTTA | GACTAATCTG | TAGCTAGAGT |
| 101 | CACTGATGAA | TCCTTTACAA | CTTTCAAAAA | AAAAA | |

FIG. IIA

| HPO4 HPO5 ATDP7 C.brew. ATOP5 S.cerev. | MLRSLLRGLT MLRSLLRGLT MSVSSLFNLP MS.SSMLNFT MTADNNSM | HIPRVNSAQQ .LIRLRSLA. .ASRIVSLPL .TGPPPRFFP | PSCAHARLQF | AHRPLSSIS. PLCFFSPISL | PRKLPNFRAF TQRFSAKLTF |
|---|--|--|--|--|--|
| | SEDRTDHMRG SGTA.MTD SSQATT.MGE T.MTD | ASTWAGGQSQ ASTWAGGQSQ TKDAGMDAVQ VVDAGMDAVQ SNDAGMDAVQ ETCFSGHDEE | DELMLKDECI RRLMFEDECI RRLMFEDECI RRLMFEDECI | LVDETDRVVG LVDENDKVVG LVDENNRVVG | HASKLECHKF HVSKYNCHLM HESKYNCHLM HDTKYNCHLM |
| | LPHOPAGLLH ENIEAKNLLH ENIESENLLH EKIEAENLLH | RAFSVFLFDD RAFSVFLFNS RAFSVFLFNS | QGRLLLQQRA KYELLLQQRS KYELLLQQRS KYELLLQQRS | RSKITFPSVW RSKITFPSVW NTKVTFPLVW ATKVTFPLVW KTKVTFPLVW TEKITFPDLW | TNTCCSHPLY TNTCCSHPLY TNTCCSHPLY TNTCCSHPLY |
| | RERE | SQVADGTVPG SELIQDNALG SELIDENCLG SELIEENVLG | AKAAAIRKLE VRNAAQRKLL VRNAAQRKLL VRNAAQRKLF | HELGIPAHOL HELGIPAHOL DELGIVAEDV DELGIPAEDL DELGIVAEDV HELGIPEDET | PA.SAFRFLT PV.DEFTPLG PV.DQFIPLS PV.DEFTPLG |
| | 201 RLHYCAADVQ RLHYCAADVQ RMLY RILY RMLY | .KAPSDGKWG .KAPSDGKWG .KAPSDGKWG | EHEMDYILFI EHEMDYILFI EHELDYLLFI EHELDYLLFI EHEVDYLLFI EHEIDYILFY | IRDVNL | APNPDEVDEV QPNPDEVAEI DPNPDEVAEV QPNPDEVAEI |

÷:

14/45

FIG. IIB

300
RYVTQEELRQ MMQ...PDN GLQWSPWFRI IAARFLERWW ADLDAALNTD
RYVTQEELRQ MMQ...PDN GLQWSPWFRI IAARFLERWW ADLDAALNTD
KYVSREELKE LVKKADAGEE GLKLSPWFRL VVDNFLMKWW DHVEKGTLVE
KYMNRDDLKE LLRKADAEEE GVKLSPWFRL VVDNFLFKWW DHVEKGSLKD
KYVSREELKE LVKKADAGDE AVKLSPWFRL VVDNFLMKWW DHVEKGTITE
KWVSPNDLKT MF....ADP SYKFTPWFKI ICENYLFNWW EQLDDLSEVE

301
KHEDWGTVHH INEA*
KHEDWGTVHH INEA*
A.IDMKTIHK L*
A.ADMKTIHK L*
A.ADMKTIHK L*
A.ADMKTIHK L*
NDRQ...IHR ML*

F1G. 12

| 1 | ccaaaaacaa | ctcaaatctc | ctccgtcgct | cttactccgc | catgggtgac |
|-----|------------|------------|------------|------------|------------|
| 51 | | tggatgctgt | | | |
| 101 | | gatgagtgtg | | | |
| 151 | | gatggagaag | | | |
| 201 | | ttctattcaa | | | |
| 251 | | aaggtgacat | | | |
| 301 | | ctacagagaa | | | |
| 351 | | gaggaxxxxx | | | |
| 401 | | ×××××××× | | | |
| 451 | | xxxxxxxxx | | | |
| 501 | | xxxxxxxxx | | | |
| 551 | | xxxxxxxxx | | | |
| 601 | | xxxxxxxxx | | | |
| 651 | | xxxxxxxxx | | | |
| 701 | | caatttgata | | | |
| 751 | | ccgaaaagca | | | |
| 801 | | tttttttt | | | |
| 851 | | tcttttgtgt | | | |
| 901 | | ttgatggttt | | | |
| 951 | atctaaaaa | | 3 | 7 | |

| 1 MECVGARNFA AMAVSTFPSW SCRRKFPVVK RYSYRNÍRFG LCSVRASGGG SSGSESCVAV REDFADEXDF | Cyanobacterial enzyme begins ———————————————————————————————————— | Dinucleotide-binding signature | 210 PKLIWPNN YGVWVDEFEA MDLLDCLDaT WSGa-VYiDd -t-KDL-RPY GRVNRKQLKS KMMQKCI-NG DLPFTNN YGVWEDEFND LGLQKCIEHV WRETIVYLDD DKPITIGRAY GRVSRRLHE ELLRRCVESG PNN YGVW-DEFLC WVY-DDR-Y GRV-RL | |
|---|--|-------------------------------------|--|---|
| kfg | GLAVAQQV GLALAAES GLA-A | ide-bindi | GRVNRKOL GRVSRRLL GRV-RL | |
| vkS-f-s- LCSVRASGGG | DLAVVGGGPA DHVVIGCGPA DV-G-GPA | Dinucleot | -t-KOL-RPY OKPITIGRAY R-Y | |
| -HGF- RYSYRNIRFG | begins —S.Kg-VV ppiSiGDGAL | | WSGa-VYiDd WRETIVYLDD WVY-DD | |
| PN-Laf1-p- SCRRKFPVVK F | Cyanobacterial enzyme begins LVPETKKKNL DFELPMYDD: S.Kg-\ FVQMQQNKDM DEQSKLVDKL PPISIGDG-\ -V | tion domain | MDLLDCLDaT LGLQKCIEHV LC | |
| MDTLLKT AMAVSTFPSW T | Yanobacter LVPETKKKNL FVQMQQNKDM -V | nit interac | YGVWVDEFEA YGVWEDEFND YGVW-DEF | |
| 1 MECVGARNFA | 71 VKSSALLa VKAGGSEIL. VKSL- | Possible subunit interaction domain | 141 -PKLIWPNN DLPFTNN |) |
| Plant <i>beto</i> A.t.epsilon Consensus | Plant <i>beta</i> A.t.epsilon Consensus | P | Plant <i>beta</i> A.t.epsilon Consensus | |

| 280 SKVDS ITKASDGLRL VACDDNNVIP CRLATVASGA ASGKLLÖYEV GGPRVCVÖTA YGVEVEVENS -KV |
|---|
| PYnPGY. QVA GGPRVCVQTA Q-A |
| SRLVQYDK ASGKLLQYEV L-QY |
| AtVVLDATGF CRLATVASGA A-G- |
| VACDDNNVIP |
| ViHE.E-KSm ITKASDGLRL |
| 211 VKFHqaKVik VSYLSSKVDS VKV |
| Plant beta A.t.epsilon Consensus |

Conserved region #1

F1G. 13 A

Predicted TM helix 13B Conserved region #3 Conserved region #4 Conserved region #5 Predicted TM helix Conserved region #2 Plant beta A.t.epsilon Plant beta A. t.epsilon Plant beta 4. t.epsilon Plant beta A.t.epsilon Consensus Consensus Consensus Consensus

FIG. 14A

Adonis palaestina ε-cyclase cDNA #5 Length: 1898

| 1 | and and the state of the state |
|------|--|
| 51 | |
| 101 | |
| | guyuducuyy cualyyddil dilliddinii cacaacetca tetettette |
| 151 | coolegigues activitique caadaaacet tantanttes assetsenti |
| 201 | acadeataca tegatatuut tetterinta nantanatti teesetasaa |
| 251 | 9009009909 GARAGUCUUUAU LANAANTTOT OTTOCTTATA 3303000+++ |
| 301 | - cycygacyda ydyydlllid llaaantan taattetaaa c++++ |
| 351 | cocadatyca ycaadcaddu liitainnana aacaggccaa actagaara |
| 401 | aagttgccac caataccttt tggagaatcc gtgatggact tggttgtaat |
| 451 | aggttgtgga cctgctggtc tttcactggc tgcagaagct gctaagctag |
| 501 | ggttgaaagt tggccttatt ggtcctgatc ttccttttac aaataattat |
| 551 | ggtgtgtggg aagacgagtt caaagatctt ggacttgaac gttgtatcga |
| 601 | gcatgettag aaggacacca tegtatatat |
| 651 | gcatgcttgg aaggacacca tcgtatatct tgataatgat gctcctgtcc |
| 701 | ttattggtcg tgcatatgga cgagttagtc gacatttgct acatgaggag |
| 751 | ttgctgaaaa ggtgtgtgga gtcaggtgta tcatatctgg attctaaagt |
| 801 | ggaaaggatc actgaagctg gtgatggcca tagccttgta gtttgtgaaa |
| 851 | digagater talectific addettacts contracts tagages act |
| 901 | tedayagaac tittigaadta Idaadtaddt doccetedta tttatatataa |
| 951 | dacegerial gyddiddadd T.T.Daddfora daacaateca tacgatecaa |
| | decidatyyi dilediydae tacadadaet atatocaaca gaaattacaa |
| 1001 | - 1901099009 0090010100 dacattrott tatotrator contators |
| 1051 | adeadyacti tittilidadd adacetafff aaceteaaa aataceataa |
| 1101 | carregates designed adactigation carrattings goetstaget |
| 1151 | diccadgita tadaddilia idaadaddaa footcatata ttoctottoo |
| 1201 | - 19911CILLA CCADACACAD ACCAAAAAAA CCTAGCATTT GGTGGTGGA |
| 1251 | - cuayearyyr yearceadea acadderatt contrataca atcactata |
| 1301 | gadgeteedd daidigelle lataattaca aagattttaa agaaagaaa |
| 1351 | |
| 1401 | aagcatggag cagtctttgg ccaaaggagc gaaaacgtca aagagcatTc |
| 1451 | tttcttttTg gattagagct tattgtgcag ctagatattg aagcaaccag |
| 1501 | aacattettt agaacettet teegettgee aacttggatg tggtggggtt |
| 1551 | tecttaggte tteactates tetttegate tectatell tagging taggaggtt |
| 1601 | tccttgggtc ttcactatca tctttcgatc tcgtcttgtt ttccatgtac |
| 1651 | atgtttgttt tggcgccaaa cagcatgagg atgtcacttg tgagacattt |
| 1701 | Secretaria Control of Contrattor sanagettae et en |
| 1751 | agreered arranged to the carrage at a second and a second arranged to the carrage at a second arranged |
| 1801 | - TOUS AND TOUR LANGUAGE AND THE CONTRACT OF THE TOUR PROPERTY OF THE TO |
| 1851 | Laddyladdi UCCUUUTTTA ATATTATTAT Atcaaacca caast l |
| 1001 | aaagtaaatt tattgataca aaagtagttt ttttccttaa aaaaaaaa |
| | |

FIG. 14B

Adonis palaestina E-cyclase #5 predicted polypeptide TRANSLATE from: 113 to: 1702 Length: 529 amino acids

MELLGVRNLI SSCPVWTFGT RNLSSSKLAY NIHRYGSSCR VDFOVRADGG SGSRSSVAYK EGFVDEEDFI KAGGSELLFV QMQQTKSMEK QAKLADKLPP 51 101 IPFGESVMDL VVIGCGPAGL SLAAEAAKLG LKVGLIGPDL PFTNNYGVWE DEFKDLGLER CIEHAWKDTI VYLDNDAPVL IGRAYGRVSR HLLHEELLKR 151 201 CVESGVSYLD SKVERITEAG DGHSLVVCEN EIFIPCRLAT VASGAASGKL LEYEVGGPRV CVQTAYGVEV EVENNPYDPN LMVFMDYRDY MOOKLOCSEE 251 301 EYPTFLYVMP MSPTRLFFEE TCLASKDAMP FDLLKRKLMS RLKTLGIOVT KVYEEEWSYI PVGGSLPNTE QKNLAFGAAA SMVHPATGYS VVRSLSEAPK 351 YASVIAKILK QDNSAYVVSG QSSAVNISMQ AWSSLWPKER KRQRAFFLFG 401 LELIVOLDIE ATRTÉFRTFF RLPTWMWGF LGSSLSSFDL VLFSMYMFVL 451 501 APNSMRMSLV RHLLSDPSGA VMVRAYLER*

- -

20/45

FIG. 15A

DNA sequence of potato cDNA (GenBank R27545) obtained from Nicholas J. Provart

```
potato.seq Length: 1378 August 2, 1996 13:06 Type: N Check: 605
     tagcggnnnn naggatgagt tcaaagatct tggtcttcaa gcctgcattg
     aacatgtttg gcgggatacc attgtatatc ttgatgatga tgatcctatt
  51
     cttattggcc gtgcctatgg aagagttagt cgccatttac tgcacgagga
 101
 151
     gttactcaaa aggtgtgtgg aggcaggtgt tttgtatcta aactcgaaag
     tggataggat tgttgaggcc acaaatggcc acagtcttgt agagtgcgag
 201
 251
     ggtgatgttg tgattccctg caggtttgtg actgttgcat cgggagcagc
 301
     ctcggggaaa ttcttgcagt atgagttggg aggtcctaga gtttctgttc
 351
     aaacagctta tggagtggaa gttgaggtcg ataacaatcc atttgacccg
 401
     agcctgatgg ttttcatgga ttatagagac tatgtcagac acgacgctca
     atctttagaa gctaaatatc caacatttct ctatgccatg cccatgtctc
 451
 501
     caacacgagt ctttttcgag gaaacttgtt tggcttcaaa agatgcaatg
     ccattcgatc tgttaaagaa aaaattgatg ttacgattga acaccctcgg
 551
     tgtaagaatt aaagaaattt atgaggagga atggtcttac ataccagttg
601
651<sup>-</sup>
     gaggatettt gecaaataea gaacaaaaaa eacttgeatt tggtgetget
 701
     gctagcatgg ttcatccagc cacaggttat tcagtcgtca gatcactgtc
 751
     tgaagctcca aaatgcgcct tcgtgcttgc aaatatatta cgacaaaatc
     atagcaagaa tatgcttact agttcaagta ccccgagtat ttcaactcaa
801
     gcttggaaca ctctttggcc acaagaacga aaacgacaaa gatcgttttt
851
901
     cctatttgga ctggctctga tattgcagct ggatattgag gggataaggt
 951
     catttttccg cgcgttcttc cgtgtgccaa aatggatgtg gcagggattt
     cttggttcaa gtctttcttn agcagacctc atgttatttg ccttctacat
1001
     gtttattatt gcaccaaatg acatgagaag aggcttaatc agacatcttt
1051
1101
     tatctgatcc tactggtgca acattgataa gaacttatct tacattttag
1151
     agtaaattcc tcctacaata gttgttgaan nagaggcctc attacttcag
1201
     attcataaca gaaatcgcgg tctctcgagg ccttgtatat aacattttca
1251
     ctaggttaat attgcttgaa taagttgcac agtttcagtt tttgtatctg
1301
     cttcttttt gtccaagatc atgtattgan ccaatttata tacattgcca
1351
     gtatatataa attttataaa aaaaaaaa
```

poteps.pep Length: 378 TRANSLATE from: 14 to: 1147

- 1 DEFKDLGLQÄ CIEHVWRDTI VYLDDDDPIL IGRAYGRVSR HLLHEELLKR
 51 CVEAGVLYLN SKVDRIVEAT NGHSLVECEG DVVIPCRFVT VASGAASGKF
 101 LQYELGGPRV SVQTAYGVEV EVDNNPFDPS LMVFMDYRDY VRHDAQSLEA
 151 KYPTFLYAMP MSPTRVFFEE TCLASKDAMP FDLLKKKLML RLNTLGVRIK
 201 ELVEEDISVI DVCCSLDNTE OVTLAEGAAA SMVUDATGVS VVDSLSGADV
- 201 EIYEEEWSYI PYGGSLPNTE QKTLAFGAAA SMVHPATGYS VVRSLSEAPK 251 CAFVLANILR QNHSKNMLTS SSTPSISTQA WNTLWPQERK RQRSFFLFGL
- 301 ALILQLDIEG ÎRSFFRAFFR VPKWMWQGFL GSSLSXADLM LFAFYMFIIA
- 351 PNDMRRGLIR HLLSDPTGAT LIRTYLTF*

FIG. 15B

Chimeric lettuce/potato lycopene $\epsilon\text{-cyclase}$: converts lycopene to $\delta\text{-}$ carotene, the lettuce cDNA converts lycopene to $\epsilon\text{-carotene}$ and the potato cDNA does not produce an active enzyme

(amino acids in lower case are from lettuce and those in uppercase are from the potato cDNA; an $Ava\Pi$ site in common to the two cDNAs was used to construct the chimera)

| 1 | mecfgarnmt | atmavftcpt | ftdcnirhkf | sllkqrrftn | lsassslrqi |
|-----|------------|-------------------|------------|------------|------------|
| 51 | kcsaksdrcv | vdkqgisvac | eedyvkaggs | elffvqmqrt | ksmesqskls |
| 101 | eklagipign | cildlvvigc | gpaglalaae | saklglnvgl | igpdlpftnn |
| 151 | ygvwqdefig | lglegciehs | wkdtlvyldd | adpirigray | grvhrdllhe |
| 201 | ellrrcvesq | vsylsskver | iteapngysl | iecegnitip | crlatvasga |
| 251 | asgkflevel | gGPRVSVQTA | YGVEVEVDNN | PFDPSLMVFM | DYRDYVRHDA |
| 301 | QSLEAKYPTF | LYAMPMSPTR | VFFEETCLAS | KDAMPFDLLK | KKLMLRLNTL |
| 351 | GVRIKEIYEE | EWSYIPVGGS | LPNTEQKTLA | FGAAASMVHP | ATGYSVVRSL |
| 401 | SEAPKCAFVL | ANILRONHSK | NMLTSSSTPS | ISTQAWNTLW | PQERKRQRSF |
| 451 | FLFGLALILQ | LDIEGIRSFF | RAFFRVPKWM | WQGFLGSSLS | XADLMLFAFY |
| 501 | METTAPNDMR | RGLIRHLLSD | PTGATLIRTY | LTF* | |

FIG. 16

| blosi | | 912 003 |
|-------|--|------------|
| Match | Ratio: 3.929 Gaps: Percent Similarity: 79.893 Percent Identity: 76. display thresholds for the alignment(s): = IDENTITY := 2 . = 1 | 1 139 |
| 151 | | |
| 1 | . DEFKDLĠĹQAĊĬĖHVWRDTĪVŸĹDDDDPĪLĪĠŔAŸĠŔVŚŔHĹĹHĖĖĹĹK | 49 |
| 201 | RCVESGVSYLSSKVDSITEASDGLRLVACDDNNVIPCRLATVASGAASGK | 250 |
| 50 | . . : . | 99 |
| 251 | LLQYEVGGPRVCVQTAYGVEVEVENSPYDPDQMVFMDYRDYTNEKVRSLE | 300 |
| 100 | . | 149 |
| 301 | | 350 |
| 150 | . . | 199 |
| 351 | LKTYEEEWSYIPVGGSLPNTEQKNLAFGAAASMVHPATGYSVVRSLSEAP | 400 |
| 200 | . | 249 |
| 401 | | 445 |
| 250 | : : . | 299 |
| 446 | | 495 |
| 300 | . | 349 |
| 496 | | |
| 350 | . .: : : : : APNDMRRGLIRHLLSDPTGATLIRTYLTF 378 | |

FIG. 17A

Adonis palaestina Ipil attcatcttc agcagcgctg tcgtactctt tctatatctt cttccatcac 51 taacagtagt cgccgacggt tgaatcggct attcgcctca acgtcaacta 101 tgggtgaagt cactgatgct ggaatggatg ctgttcagaa gcggctcatg ttcgacgacg aatgtatttt ggtggatgag aatgacaagg tcgtcgggca 151 201 tgattccaaa tacaactgtc atttgatgga aaagatagag gcagaaaatt 251 tgcttcacag agccttcagt gttttcttgt tcaactcaaa atatgaattg 301 cttcttcagc aacgatccgc cacaaaggta acattcccgc tcgtatggac 351 aaacacatgt tgcagtcatc ctctctttcg tgattccgag ctcatagaag 401 aaaattatct cggtgtacga aacgctgcac aaagaaagct tttagacgag 451 ctaggcattc cagctgaaga tgtcccagtt gatgaattta ctcctcttgg 501 tcqcattctt tacaaagctc catctgacgg caaatgggga gagcacgaat 551 tggactatct cctatttatt gtccgagatg tgaaatacga tccaaaccca 601 gatgaagttg ctgatgctaa gtatgttaat cgcgaggagt tgagagagat 651 actgagaaaa gctgatgctg gtgaagaggg actcaagttg tctccttggt ttagattggt tgttgataac tttttgttca agtggtggga tcatgtagag 701 751 cagggtacga ttaaggaagt tgctgacatg aaaactatcc acaagttgac 801 ttaagaggac ttctctcctc tgttctacta tttgtttttt gctacaataa gtgggtggtg ataagcagtt tttctgtttt ctttaattta tggcttttga 851 atttgcctcg atgttgaact tgtaacatat ttagacaaat atgagacctt 901 951 gtaagttgaa tttgaggctg aatttatatt tttgggaaca taataatgtt 1001 aa

÷:

24/45

FIG. 17B

Adonis palaestina Ipi2 ttttaaagct ctttcgctcc accaccatca aagccagcca aatttctctg tacaaaagtt aaaaacaccg ctttgggctt tggcccctcc atatcggaat 51 ccttgtttac gatacgcatc taaaccagta attctcggtt ttaatttgtt 101 tcctaaatta ggcccctttc cggaatcccg agaattatgt cgtcgatcag 151 gattaatcct ttatatagta tcttctccac caccactaaa acattatcag 201 cttcgtgttc ttctcccgct gttcatcttc agcagcgttg tcgtactctt 251 tctatttctt cttccatcac taacagtcct cgccgagggt tgaatcggct 301 gttcgcctca acgtcgacta tgggtgaagt cgctgatgct ggtatggatg 351 ccqtccaqaa qcgqcttatg ttcgacgatg aatgtatttt ggtggatgag 401 aatgacaagg tcgtcggaca tgattccaaa tacaactgtc atttgatgga 451 aaagatagag gcagaaaact tgcttcacag agccttcagt gttttcttat 501 tcaactcaaa atacgagttg cttcttcagc aacgatctgc aacgaaggta 551 acattecege tegtatggae aaacacetgt tgeagecate ecetetteeg 601 tgattccgaa ctcatagaag aaaattttct cggggtacga aacgctgcac 651 aaaggaagct tttagacgag ctaggcattc cagctgaaga cgtaccagtt 701 gatgaattca ctcctcttgg tcgcattctt tacaaagctc catctgacgg 751 aaaatgggga gagcacgaac tggactatct tctgtttatt gtccgagatg 801 tgaaatacga tccaaaccca gatgaagttg ctgacgctaa gtacgttaat 851 cgcgaggagt tgaaagagat actgagaaaa gctgatgcag gtgaagaggg 901 aataaagttg tctccttggt ttagattggt tgtggataac tttttgttca 951 agtggtggga tcatgtagag gaggggaaga ttaaggacgt cgccgacatg 1001 aaaactatcc acaagttgac ttaagagaaa gtctcttaag ttctactatt 1051 tggtttttgc ttcaataagt ggatggtgat gagcagtttt tatgcttcct 1101 ttaattttgg cttttcaatt tgctttatgt gttgaacttg taacatattt 1151 agtcaaatat gagaccttgt gagttgaatt tgaggttata tttatagttt 1201 tgggaacata aaaaaaaaaa 1251

FIG. 18A

Haematococcus pluvialis Ipil 1 ctcggtagct ggccacaatc gctatttgga acctggcccg gcggcagtcc gatgccgcga tgcttcgttc gttgctcaga ggcctcacgc atatcccccg 51 cgtgaactcc gcccagcagc ccagctgtgc acacgcgcga ctccagttta 101 agctcaggag catgcagatg acgctcatgc agcccagcat ctcagccaat 151 ctgtcgcgcg ccgaggaccg cacagaccac atgaggggtg caagcacctg 201 ggcaggcggg cagtcgcagg atgagctgat gctgaaggac gagtgcatct 251 tggtggatgt tgaggacaac atcacaggcc atgccagcaa gctggagtgt 301 cacaagttcc taccacatca gcctgcaggc ctgctgcacc gggccttctc 351 tgtgttcctg tttgacgatc aggggcgact gctgctgcaa cagcgtgcac 401 gctcaaaaat caccttccca agtgtgtgga cgaacacctg ctgcagccac 451 cctttacatg ggcagacccc agatgaggtg gaccaactaa gccaggtggc 501 cgacggaaca gtacctggcg caaaggctgc tgccatccgc aagttggagc 551 acgagetggg gataccageg caccagetge eggeaagege gtttegette 601 ctcacgcgtt tgcactactg tgccgcggac gtgcagccag ctgcgacaca 651 atcagcgctc tggggcgagc acgaaatgga ctacatcttg ttcatccggg 701 ccaacgtcac cttggcgccc aaccctgacg aggtggacga agtcaggtac 751 gtgacgcaag aggagctgcg gcagatgatg cagccggaca acgggctgca 801 atggtcgccg tggtttcgca tcatcgccgc gcgcttcctt gagcgttggt 851 gggctgacct ggacgcggcc ctaaacactg acaaacacga ggattgggga 901 acggtgcatc acatcaacga agcgtgaaag cagaagctgc aggatgtgaa 951 gacacgtcat ggggtggaat tgcgtacttg gcagcttcgt atctcctttt 1001 tctgagactg aacctgcagt caggtcccac aaggtcaggt aaaatggctc 1051 gataaaatgt accgtcactt tttgtcgcgt atactgaact ccaagaggtc 1101 aaaaaaaaa aaaaa 1151

J. 12

26/45

FIG. 18B

Haematococcus pluvialis Ipi2

```
tggaacctgg cccggcggca gtccgatgcc gcgatgcttc gttcgttgct
 51
     cagaggeete acgeatatee egegegtgaa eteegeecag cageecaget
101
     gtgcacacgc gcgactccag tttaagctca ggagcatgca gctgcttgcc
151
     gaggaccgca cagaccacat gaggggtgca agcacctggg caggcgggca
201
     gtcgcaggat gagctgatgc tgaaggacga gtgcatctta gtggatgctg
251
     acgacaacat cacaggccat gccagcaagc tggagtgcca caaattccta
301
     ccacatcage etgeaggeet getgeacegg geettetetg tgtteetgtt
351
     tgacgaccag gggcgactgc tgctgcaaca gcgtgcacgc tcaaaaatca
401
     ccttcccaag tgtgtggacg aacacctgct gcagccaccc tctacatggg
451
     cagaccccag atgaggtgga ccaactaagc caggtggccg acggcacagt
501
     acctggcgca aaagctgctg ccatccgcaa gttggagcac gagctgggga
551
     taccagcgca ccagctgccg gcaagcgcgt ttcgcttcct cacgcgtttg
601
     cactactgtg ccgcggacgt gcagccggct gcgacacaat cagcgctctg
651
      gggcgagcac gagatggact acatcttatt catccgggcc aacgtcacct
701
      tggcgcccaa ccctgacgag gtggacgaag tcaggtacgt gacgcaagag
751
     gagctgcggc agatgatgca gccggacaac gggttgcaat ggtcgccgtg
     gtttcgcatc atcgccgcgc gcttccttga gcgttggtgg gctgacctgg
801
851
      acgcggccct aaacactgac aaacacgagg attggggaac ggtgcatcac
     atcaacgaag cgtgaaggca gaagctgcag gatgtgaaga cacgtcatgg
901
 951
      ggtggaattg cgtacttggc agcttcgtat ctcctttttc tgagactgaa
1001
      cctgcagagc tagagtcaat ggtgcatcat attcatcgtc tctcttttgt
1051
      tttagactaa tctgtagcta gagtcactga tgaatccttt acaactttca
1101
      aaaaaaaa
```

...

F1G. 19A

Lactuca sativa Ipil tgccaaaatg ttgaaatttc ccccttttaa aaccattgct accatgatct cttctccata ttcttccttc ttgctgcctc ggaaatcttc tttccctcca atgccgtctc tcgcagccgc tagtgttttc ctccaccctc tttcgtctgc 101 cgctatgggc gattccagca tggatgctgt ccagcgacgt ctcatgttcg 151 atgacgaatg cattitiggtg gatgagaatg acaaagtggt tggccatgat actaaataca attgtcattt gatggagaag attgaaaagg gaaatatgct 201 251 acacagagca ttcagtgtgt tcttgttcaa ctcgaaatat gaattactcc 301 ttcagcaacg ttctgcaacc aaggtgactt tccctttggt atggacaaac acgtgttgca gccatccact atacagggag agtgagctta ttgacgaaaa 351 401 cgcccttggg gtgaggaatg ctgcacagag gaagctcctg gatgaactcg gcatcctgg agcagatgtt ccggttgatg agttcactcc attgggtcgc 451 501 ăttctatacă ağgccgcătc ggatggaaag tggggagaac atgaacttga 551 ttacctgctg tffatggtac gfgatgttgg tffggatccg aacccagatg 601 aagtgaaaga tgtaaaatat gtgaaccggg aagagctgaa ggaattggta aggaaggcgg atgctggtga agagggtgtg aagctgtccc cgtggttcaa attgattgtc gataatttct tgtttcagtg gtgggatcga ctccataagg 651 701 751 801 gaaccctaac cgaagctatt gatatgaaaa caatccacaa actcacataa ăaacactaca ctagtaggag ăgaggăttat atgagatatt tgttatatgt 851 gaaattgaaa ttcăgatgaă tgcttgtatt tatttctatt tggacaaact 901 tcaacttctt tttgctacct tatcagaaaa aaaaa 951

FIG. 19B

Lactuca sativa Ipi2 tattcgcttc aaaatctctt ccattaactg ctcaaatctc caccttcgcc ggtctťaatc tccgccggcg cactttcacc accataaccg ccgccatggg tgacgattcc ggcatggacg ctgtccagag acgtctcatg tttgatgatg 101 aătgčatttt ğğttgătgaă aatgacaatg ttcttgggca tgataccaaa 151 tacăattgtc ăcttgatgga gaagattgag aaagatăătt tgcttcatag agcattcagt gtattttat tcaattcaaa atacgaatta ctccttcagc 201 251 aaaggtcaga aaccaaggtg acatttcctt tggtatggac aaacacctgt 301 tgcăğccatc cactatăcağ agaatcggag ttaattcccg aaaatgccct tggggtcaga aatgctgcac agaggaagct tctagatgaa ctcggtatcc 351 401 ctgctgaaga tgttccagtt gatgagttca caactttagg tcgcatgttg 451 tacaaggete catetgatgg aaaatggggt gaacatgaag ttgattacet 501 acteticete gtgcgtgacg ttgccgtgaa cccaaaccet gatgaggtgg 551 cggacattag ătăcğtgaac caagaagat taaaagagtt actaaggaag 601 gcggatgcgg gtgaggggg tttgaaattg tccccatggt ttaggctagt 651 ğgtggacaac ttcttgttca aatggtggga tcatgtccaa aaggggacac 701 tcaatgaagc aattgacatg aaaaccattc ataagttgat atgaaaaatg 751 gttaatatt atggtggtgg tttggagcta ataatttgtg tgttcaagtc tcggtccttc ttttttaac gtttttttt tttcttttat tgggagtgtt 801 851 tattgtgtac ttgtaacgta ggccctttgg ttacgcttta agagtttaat 901 aaagaaccac cgttaattta aaaaaaaaa aaaaaaaa

FIG. 20

Chlamydomonas reinhardtii Ipil

(Note: the isomerase cDNA probably ends at ca. base 1103; the second half of the cDNA is similar to extensin and other hydroxyproline-rich structural proteins)

| 1 51 101 151 201 251 301 351 401 451 501 551 601 651 | ggcacgagct aactacctca cgctgttttc gatgtcataa ccagagcgca caaagccgtg aggactttca cctgggaagg tgcttggtgg cgactgccac accgcgcctt cagcagcgcg ctgctgctcg cggcggcggt | tctgctccag actcccactt acctgtctta | taactcgcgg | atcgggaatt caacacattt gcattgcttt cagtttcatc gggcgtccat tccgccctgt gcgaatggcg acttcatgca ctaggcaccg ccagccctgc ccgacggccg ccgggtgtgt gccggacgag gcatcaaggc | tggaagcttg cgcgcgccat agatcgcttt gagcccaagc gcgcctcgcg ggccgggagc agttcgtcaa gcgggacgag ccaacaagta ggccgcctgc actgctgctg ggaccaacac gtggacctgc ggcggcggtg |
|---|--|--|--|--|---|
| 1101 1151 1201 1251 1301 | cggcagccga aaaggggaag ttgtgatgcg cgggcgtgag gcgatgggta gcatagcgtg | caggggcggg gcgtgggatg cgtgtgtgta catgtgtgtg | agcgggggat aggtctgaag cgtgagcgac cggagggtcg gcggctgcgc | gaatgggaat acagggggaa aaagccggga gtgggtcggt gggtatgtgg | gtgaatgcga aatcgggggg ggcggaccgc cggttgcgcg gcacccgggc |
| 1351 1401 1451 1501 | acggaggaga ggcgggcctc ggggctgcac tcacttggtg | aggcácácge actectggte ccatatgage aggtggggeg | aggtggcgcg gtgcccagtg ggcgcactgc aggtggctgt | gaggtgtgtc gtctcgtggg cgcgctgggc gggcggcggg | aggggccatg cagagtggca taagtcctta cgcagtggca |
| 1551 1601 1651 1701 1751 | gaaggacacg ggcggatagc tgcaggccgc cgttggggag gggcgcctga | gatatgacgt gagaagcggg gtgccgcctg gtagtggcgg | cggtggcagg caggcgcggc ccacaggagg | cgctgtaatg aggccgcagg gcggggggg cgcaggaggc | cgggagaatg ctgcagcacc cctgagtaat agcagcagga |
| 1801 1851 | ggacgagctg gtggccata | g gagggacccg c aaaaaaaaaa | ttggcaaccc aaaa | . uuggeege | , 30304464 |

رز جر

FIG. 21A

Tagetes erecta Ipil ccaaaaacaa ctcaaatctc ctccgtcgct cttactccgc catgggtgac gactccggca tggatgctgt tcagcgacgt ctcatgtttg acgatgaatg 101 cattttggtg gatgagtgtg acaatgtggt gggacatgat accaaataca attgtcactt gatggagaag attgaaacag gtaaaatgct gcacagagca ttcagcgttt ttctattcaa ttcaaaatac gagttacttc ttcagcaacg 151 201 251 gtctgcaacc aaggtgacat ttcctttagt atggaccaac acctgttgca ğccatccact ctăcağagaa tccgagcttg ttcccgaaaa cgcccttgga gtaagaaatg ctgcacagag gaagctgttg gatgaactcg gtatccctgc 301 351 tgaagatgtt cccgttgatc agtttactcc tttaggtcgc atgctctaca 401 aggetecate tgatggaaag tggggagaac atgaaettga etacetaett tteatagtga gagaegttge tgtaaaeeeg aacceagatg aagtggegga 451 501 551 tatcaaatat gtganccang aagagttaaa ggagctgcta aggaaagcag atgcggggga ggagggtttg aagctgtctc catggttcag gttagtggtt 601 gataacttet tgttcaagtg gtgggatcat gtgcaaaagg gtacactcac 651 tgaagcaatt gatatgaaaa ccatacacaa gctgatatag aaacacaccc 701 tčaačcgaaa agttcaagcc taataattcg ggttgggtcg ggtctaccat caattgttt tttctttaa gaagttttaa tctctatttg agcatgttga 751 801 851 ttcttgtctt ttgtgtgtaa gattttgggt ttcgtttcag ttgtaataat 901 gaaccattga tggtttgcaa tttcaagttc ctatcgacat gtagtgatct 951 aaaaaa

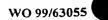
FIG. 21B

Oryza sative Ipil cctccctttg cctcgcgcag aggcggccgc gccttctccg ccgcgaggat ggccggcgcc gccgccgcg tggaggacgc cgggatggac gaggtccaga agcggctcat gttcgacgac gaatgcattt tggtggatga acaagacaat 101 gttgttggcc atgaatcaaa atataactgc catctgatgg aaaaaatcga 201 atctgaaaat ctacttcata gggctttcag tgtattcctg ttcaactcaa aatafgaact cctactccag caacgatctg caacaaaggt tacatttcct ctagtttgga ccaacacttg ctgcagccat cctctgtacc gtgagtctga 251 301 gcttatacag gaaaactacc ttggtgttag aaatgctgct cagaggaagc 351 tcttggatga gctgggcatc ccagctgaag atgtgccagt tgaccaattc acccctcttg gtcggatgct ttacaaggcc ccatctgatg gaaaatgggg 401 451 501 tgaacacgag čttgactacc tgctgttcat cgtccgcgac gtgaaggtag tcccgaaccc ggacgaagtg gccgatgtga aatacgtgag ccgtgagcag ctgaaggagc tcatccgcaa agcggacgcc ggagaggaag gcctgaagct 551 601 gtčtcččtěg ttccggčtgg tfgffgačaa čftčcfčatě ěgctěgtěgg 651 atcacgtcga gaaaggcacc ctcaacgagg ccgtggacat ggagaccatc cacaagctga agtaaggact gcgatgttgt ggctggaaag aatgatcctg 701 751 aagactctgt tcttgtgctg ctgcatatta ctcttaccag ggaagttgca 801 gaagtcagaa gaagcttttg tatgtttctg ggtttggagc ttggaagtgt tgggctctgc tgactgagag attcccttat agagtgtcta tgttaattta 851 901 gcaaacttct atattataca tgattagtta attgttcggt gtctgaataa 951 āgaacaatag catgttccat gtttatttgc t 1001

| 241 255 256 270 VAVNPNPDEVADIKY VSHEELKELLRKADA 188 VGLDPNPDEVADIKY VNREELKELLRKADA 236 VAVNPNPDEVADIRY VNREELKELLRKADA 188 VKYDPNPDEVADAKY VNREELKEILRKADA 190 VKYOPNPDEVADAKY VNREELREILRKADA 190 VKVQPNPDEVAEIKY VSREELKELVKKADA 190 VKLQPNPDEVAEIKY VSREELKELVKKADA 190 VTLAPNPDEVDEVRY VTQEELRQMMQP 247 VTLAPNPDEVDEVRY VTQEELRQMMQP 259 VSLQPNPDEVDATRY VTLPELQSMMA 259 | Tagetes erecta (marigold) Lactuca sativa (romaine lettuce) Lactuca sativa (romaine lettuce) Adonis palaestina (pheasant's eye) Adonis palaestina (pheasant's eye) Oryza sativa (rice) Arabidopsis thaliana Haematococcus pluvialis Chlamydomonas reinhardtii |
|--|---|
| 240 241 20 VAVNPNPDEV 20 VGLDPNPDEV 20 VKYDPNPDEV 20 VKYDPNPDEV 20 VKYQPNPDEV 20 VKYQPNPDEV 30 VKYQPNPDEV | Tagetes er Lactuca so Lactuca so Adonis pal Adonis pal Oryza sati Arabidopsi Arabidopsi Haematocoo Haematocoo |
| AGRKLLDELGIPAED VPVDGFTPLGRMLYKAPSDGKAG EHELDYLLFIVRD VAVNPNPDEVADJKY VSHEELKELLRKADA AQRKLLDELGIPAED VPVDGFTPLGRMLYKAPSDGKAG EHELDYLLFIVRD VGLDPNPDEVADJKY VSHEELKELLRKADA AQRKLLDELGIPAED VPVDEFTPLGRILYKAPSDGKAG EHELDYLLFIVRD VAVNPNPDEVADJRY VNREELKELLRKADA AQRKLLDELGIPAED VPVDEFTPLGRILYKAPSDGKAG EHELDYLLFIVRD VKYDPNPDEVADJRY VNREELKEILRKADA AQRKLLDELGIPAED VPVDEFTPLGRILYKAPSDGKAG EHELDYLLFIVRD VKYDPNPDEVADJAKY VNREELREILRKADA AQRKLLDELGIPAED VPVDEFTPLGRMLYKAPSDGKAG EHELDYLLFIVRD VKYOPNPDEVADJKY VSREELKELIRKADA AQRKLLDELGIVAED VPVDEFTPLGRMLYKAPSDGKAG EHELDYLLFIVRD VKVQPNPDEVADJKY VSREELKELVKKADA AJRKLEHELGIPAHQ LPASAFRFLTRLHYC AADVQPAATQSALMG EHENDYLLFIRAN VTLAPNPDEVDEVRY VTQEELRQYMQPAIRKLEHELGIPAHQ LPASAFRFLTRLHYC AADVQPAATQSALMG EHENDYLLFIRAN VTLAPNPDEVDEVRY VTQEELRQYMQPAVRKLUBELGIPAHQ LPASAFRFLTRLHYC AADVQPAATQSALMG EHENDYLLFIRAN VTLAPNPDEVDEVRY VTQEELRQYMQPAVRKLUBELGIPAHQ LPASAFRFLTRLHYC AADVQPAATQSALMG EHENDYLLFIRAN VTLAPNPDEVDEVRY VTQEELRQYMQPAVRKLQHELGIPPEQ VPASSFSFLTRLHYC AADTATHG-PAAEWG EHEVDYVLFVRPQQP VSLQPNPDEVDEVRY VTQEELRQYMA | GEEGLKLSPWFRLVV DNFLFRAWDHVQK GTLTEAIDMKTI HKLI GEEGVKLSPWFRLVV DNFLFRAWDHVQK GTLTEAIDMKTI HKLI 232 GEEGLKLSPWFRLVV DNFLFRAWDHVQK GTLNEAIDMKTI HKLT 229 GEEGLKLSPWFRLVV DNFLFRAWDHVEE GKIKDVADMKTI HKLT 234 GEEGLKLSPWFRLVV DNFLFRAWDHVER GTIKEVADMKTI HKLT 234 GEEGLKLSPWFRLVV DNFLMKWWDHVEK GTLNEAVDMETI HKLK 236 GEEGLKLSPWFRLVV DNFLMKWWDHVEK GTLVEAIDMKTI HKL 237 GEEGLKLSPWFRLVV DNFLMKWWDHVEK GTLTEAADMKTI HKL 238 -DNGLQWSPWFRIIA ARFLERWAADLDA ALNTDKHEDWGTV HHINEA 305 -DPGLSWSPWFRILA TQPAFLPAWWGDLKR RWRPGGSRLSDWGTI HRVM 307 |
| 1 T.erecta 1 2 L.sativa 1 3 L.sativa 2 4 A.palaestina 2 5 A.palaestina 1 6 O.sativa 1 7 A.thaliana 1 8 A.thaliana 2 9 H.pluvialis 1 10 H.pluvialis 2 | 1 T.erecta 1 2 L.sativa 1 3 L.sativa 2 4 A.palaestina 2 5 A.palaestina 1 6 O.sativa 1 7 A.thaliana 1 8 A.thaliana 2 9 H.pluvialis 1 10 H.pluvialis 2 11 C.reinhardtii 1 |

F16.22B

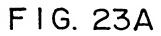
+5



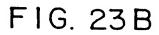
32 / 45 FIG. 24 A FIG. 23A FIG. 23 FIG. 24B FIG. 23B FIG. 24 FIG.23C FIG. 25A FIG.23D FIG. 25

FIG. 25B FIG. 25C

FIG.28A FIG. 26A FIG. 28 FIG. 26 FIG. 26B FIG. 28B



| Comparison using GAP program of the Genetics Computer Group Gap Weight: 50 Average match: 10.000 Length Weight: 3 Average Mismatch: 0.000 Quality: 17392 Length: 1904 Ratio: 9.411 Gaps: 3 Percent Similarity: 95.331 Percent Identity: 95.331 Match display thresholds for the alignment(s): = IDENTITY := 5 . = 1 |
|--|
| Adonis palaestina ϵ -cyclase #3 x Adonis palaestina ϵ -cyclase #5 |
| 1 gagagaaaaagagtgttatattaatgttactgtcgcattcttgcaacac. 49 |
| 1aaaggagtgttctattaatgttactgtcgcattcttgcaacact 44 |
| 50 .atattcagactccattttcttgttttctcttcaaaacaacaactaatg 98 |
| 45 tatattcaaactccattttctttttcttttcaaaacaaca |
| 99 tga.cggagtatctagctatggaactacttggtgttcgcaacctcatctc 147 |
| 95 tgagcagagtatetggetatggaactaettggtgttegeaaceteatete 144 |
| 148 ttcttgccctgtctggacttttggaacaagaaaccttagtagttcaaaac 197 |
| 145 ttcttgccctgtgtggacttttggaacaagaaaccttagtagttcaaaac 194 |
| 198 tagcttataacatacatcgatatggttcttcttgtagagtagattttcaa 247 |
| 195 tagettataacatacategatategeteettettettegagagtagattetteaa 244 |
| 248 gtgagggctgatggtggaagcgggagtagaacttctgttgcttataaaga 297 |
| 245 gtgágagótgátggtágáágóggágtágáagttótgttgóttátááágá 294 |
| 298 gggttttgtggacgaggaggattttatcaaagctggtggttctgagcttt 347 |
| 295 gggttttgtgaagaggattttatcaaagctggtggttctgagcttt 344 |
| 348 tgtttgtccaaatgcagcaaacaaagtctatggagaaacaggccaagctc 397 |
| 345 tgtttgtccaaatgcagcaaacaaagtctatggagaaacaggccaagctc 394 |



| 398 | gccgataagttgccaccaatacctttcggagaatctgtgatggacttggt | 447 |
|-----|--|--------------------------|
| 395 | gccgataagttgccaccaataccttttggagaatccgtgatggacttggt | |
| 448 | tgtaataggttgtggacctgctggtctttcactggctgcagaagctgcta | 497 |
| 445 | tgtaataggttgtggacctgctggtctttcactggctgcagaagctgcta | 494 |
| 498 | agctaggcttgaaagttggccttattggtcctgatcttccttttacaaat | 547 |
| 495 | agctagggttgaaagttggccttattggtcctgatcttccttttacaaat | 544 |
| | 111111111111111111111111111111111111111 | |
| | aattatggtgtgtgggaagacgagttcaaagatcttggacttgaacgttg | |
| | tatcgagcatgcttggaaggacaccatcgtatatcttgacaatgatgctc | |
| | tatcgagcatgcttggaaggacaccatcgtatatcttgataatgatgctc | |
| | ctgtccttattggtcgtgcatatggacgagttagccggcatttgctgcat | |
| 698 | gaagagttgctgaaaaggtgtgtcgagtcaggtgtatcatatctgaattc | |
| 695 | gaggagttgctgaaaaggtgtgtggagtcaggtgtatcatatctggattc | |
| 748 | taaagtggaaaggatcactgaagctggtgatggccatagtcttgtagttt | 797 |
| 745 | taaagtggaaaggatcactgaagctggtgatggccatagccttgtagttt | 794 |
| 798 | gtgaaaacgacatctttatcccttgcaggcttgctactgttgcatctgga | 847 |
| 795 | gtgaaaatgagatctttatcccttgcaggcttgctactgttgcatctgga | 844 |
| | gcagcttcagggaaacttttggagtatgaagtaggtggccctcgtgtttg | |
| | | |
| | tgtccaaactgcttatggtgtggaggttgaggtggagaacaatccatacg | 947 |
| 395 | LULCCAAACCOCTTATOOOTOOAOOTTOAOOTOOACAAFCCAFACO | $\Omega \Lambda \Lambda$ |

FIG. 23C

| 948 व | atcccaacttaatggtatttatggactacagagactatatgcaacagaaa | 997 |
|--------|--|--------------|
| 945 a | | 994 |
| 998 1 | ttacagtgctcggaagaagaatatccaacatttctctatgtcatgcccat | 1047 |
| 995 | ttacagtgctcggaagaagaatatccaacatttctctatgtcatgcccat | 1044 |
| 1048 (| gtcgccaacaagacttttttttgaggaaacctgtttggcctcaaaagatg | 1097 |
| 1045 | gtcgccaacaagacttttttttgaggaaacctgtttggcctcaaaagatg | 1094 |
| 1098 | ccatgcctttcgatctactgaagagaaaactaatgtcacgattgaagact | 1147 |
| 1095 | ccatgccattcgatctactgaagagaaaactgatgtcacgattgaagact | 1144 |
| | ctgggtatccaagttacaaaaatttatgaagaggaatggtcttatattcc | 1197 |
| | ctgggtatccaagttacaaaagtttatgaagaggaatggtcatatattcc | 1194 |
| | tgttgggggttctttaccaaacacagagcaaaagaacctagcatttggtg | 1247 |
| | tgttggtggttctttaccaaacacagagcaaaagaacctagcatttggtg | 1244 |
| | ctgcagcaagcatggtgcatccagcaacaggctattcggttgtacgatca | 1297 |
| | ctgcagcaagcatggtgcatccagcaacaggctattcggttgtacggtca | 1294 1347 |
| | ctatcagaagctccaaaatatgcttctgtaattgcaaagattttgaagca | |
| | ctgtcagaagctccaaaatatgcttctgtaattgcaaagatttgaagca | |
| | agataactctgcatatgtggtttctggacaaagcagtgcagtaaacattt | |
| | agataactctgcgtatgtggtttctggacaaagtagtgcagtaaacattt | |
| | caatgcaagcatggagcagtctttggccaaaggagcgaaaacgtcaaaga | |
| | caatgcaagcatggagcagtctttggccaaaggagcgaaaacgtcaaaga gcattctttcttttcgggttagagcttattgtgcagctagatattgaagc | • |
| | gcattctttctttcgggttagagcttattgtgcagctagatattgaagc | |
| C++1 | year terretely actually a second of the seco | |

| 1498 | aaccagaacgttctttagaaccttcttccgcttgccaacttggatgtggt | 1547 |
|------|--|------|
| 1495 | aaccagaacattctttagaaccttcttccgcttgccaacttggatgtggt | 1544 |
| 1548 | ggggtttccttgggtcttcactatcatctttcgatcttgtattgtttcc | 1597 |
| 1545 | ggggtttccttgggtcttcactatcatctttcgatctcgtcttgttttcc | 1594 |
| 1598 | atgtacatgtttgttttggccccgaacagcatgaggatgtcacttgtgag | 1647 |
| 1595 | atgtacatgtttgttttggcgccaaacagcatgaggatgtcacttgtgag | 1644 |
| 1648 | acatttgctttcagatccttctggtgcagttatggttaaagcttacctcg | 1697 |
| 1645 | acatttgctttcagatccttctggtgcagttatggtaagagcttacctcg | 1694 |
| | aaaggtaatctgttttatgaaactatagtgtctcattaaataaatga | 1744 |
| | aaaggtagtctcatctattattaaactctagtgtttcaccaaataaat | 1744 |
| | ggatccttcgtatatgtatatgatcatctctatgtatatcctatattcta | 1794 |
| | ġġatċċttċġaatgtġtatatġatċatċtċtatġtatatċċtgtactċta | 1794 |
| | atctcataaagtaatcgaaaattcattgatagaaaaaaaa | 1844 |
| | àtctcataaagtaaatgccgggtttgatattgttgtgtcaaaccggccaa | 1844 |
| | aaaa | 1848 |
| 1845 | tgatataaagtaaatttattgatacaaaagtagtttttttt | 1894 |

FIG. 23D

| • | | 0. | | | | |
|--------|---|--|---|-------------------------|---------------------------------------|---|
| | ogram of Ge 62.cmp | netics Compu | ter Group | FIG. | 24 A | : |
| Perce | Gap Weigh Length Weigh Quali Rat nt Similari display thr | ht: 4 ty: 2728 io: 5,147 ty: 99,623 | Average N | Length: Gaps: Identity: | 2.912 -2.003 530 0 99.057 | |
| Adonis | palaestina | ε-cyclase | #3 x <i>Adonis</i> , | palaestina | ε-cyclase # | 5 |
| | | | SSKLAYNIHRYO SSKLAYNIHRYO | | | |
| | 111.11111111 | 1111111111 | SELLFVQMQQTK SELLFVQMQQTK | | | |
| 1 | | 11111111111 | EAAKLGLKVGLI EAAKLGLKVGLI | | 1111 | |
| | | | NDAPVLIGRAYO NDAPVLIGRAYO | | | |
| | | | LVVCENDIFIPO | | | |
| 1 | | | NPYDPNLMVFME | | | |
| | | 11111111111 | SKDAMPFDLLKF SKDAMPFDLLKF | | | |

| 351 | KIYEEEWSYIPVGGSLPNTEQKNLAFGAAASMVHPATGYSVVRSLSEAPK | 400 |
|-----|--|-----|
| 351 | : | 400 |
| 401 | YASVIAKILKODNSAYVVSGOSSAVNISMOAWSSLWPKERKRORAFFLFG | 450 |
| 401 | | 450 |
| 451 | LELIVOLDIEATRTFFRTFRLPTWMWWGFLGSSLSSFDLVLFSMYMFVL | 500 |
| 451 | | 500 |
| 501 | APNSMRMSLVRHLLSDPSGAVMVKAYLER* 530 | |
| 501 | | |

FIG. 24B



| | | a L |
|-------------|---|--|
| | . 103 102 102 100 100 100 100 100 100 100 100 | 54 205 204 204 209 202 202 191 181 182 182 183 |
| ** | | |
| | | 160 * 180 * 200 * |
| - | KILVE KILVE KILSE KILSE TIKKE TIKKE | |
| 100 | HEKOK HEKOK HEKOK HESOS BOAGS BOAGS LIVPE LIVPE LIVPE LIVPE | **===================================== |
| | ALLE ALLE ALLE ALLE ALLE ALLE ALLE | WANTER BEAUTIES OF THE SERVICE OF TH |
| * | COMOCO COMO | 24 A A C C C C C C C C C C C C C C C C C |
| | ILFV LLFV LLFV VCVK VCVK VCVK | NEST STREET OF THE STREET OF T |
| | GGSE GGSE GGSE GGSE GGSE GGSE GGSE GGSE | DDDP DDKP VDAP VDAP DDCP DNDP NNTK NNTK KSTK KSTK |
| 8 | DEVKE DETKE DETKE DETKE DETKE DETKE TIL TIL TIL TIL TIC TIC TIC TIC | |
| | ADEER VDEER VDEER VDEER VDEER VACE VFCE VFCE VFCE VFCE VFCE VFCE VFCE VF | RETT RETT ROTT ROTT ROTY ROTY SGAA SGAA SGAA SGAA |
| * | OF GF GF GF CGWF- CGWF CGW | 180 TEHVO TEHVO TO DATE |
| | AVRE AYKE AYKE VUKQ VUKQ KLQN KCQK KSQF KKHK | |
| 9 | SRSSV SRTSV SRTSV SRTSV SRTSV SDSCV NSTS TFRSV TFRSV TFRSV SPSP | KOLG KOLG KOLG KOLG KOLG KOLG KOLG KOLG |
| | 35SG 36SG 36SG 36SG 36SG 36SG 36SG 36SG 36 | |
| | ASGG VRAD VRAD VRAD VRAD VRAD SCN SKCN SKCN SKCN SKCN SKCN SKCN SKCN | YGVW YGVW YGVW YGVW |
| * | VOFQ VOFQ VOFG VE VHGF VHGF SNKF | NNAW NNAW NNAW NNAW NNAW NNAW NNAW NNAW |
| | YRNI RFGL - CSV RASGGGSSGS ESCVAVREDF ADEEDFYKAGGSE I LFVQMQQNKDMDEQSKLVDKIJPP IS YNI HRYGSSCRVDFQVRADGGSGSRSSVAYKEGF VDEEDFIKAGGSEL LFVQMQQTKSMEKQAKLADKIJPP IP YNI HRYGSSCRVDFQVRADGGSGSRTSVAYKEGF VDEEDFIKAGGSEL LFVQMQQTKSMEKQAKLADKIJPP IP RRFTNL SA - SSSLRQI KCSAKSDR CVVDKQG I SVADEEDFYKAGGSEL LFVQMQQTKSMEKQAKLADKIJP IP EKS I FLAY - EQY ESKCNSSSGSDSCVVDKEDF EEEDFYKAGGSEL LFVQMQQTKSMESQSKL SEKIJAQI P KTPNKLDF I PQFHGFE LCSNNPYHSRVRLGVKRA I KIV SSVVSGSAAL LDL VPETKKENIJ DFEL RTHNKL BFL PTLHGFA EKQHL VSTSKLQNQYFRIASRNIH PCRNGTVKASSSAL LEL VPETKKENIJ DFEL KTPNNL EFL LPPHHGF AVKAS FRSEKHHNFGSRKFCEGLG KGVCVKASSSAL LEL VPETKKENIJ DFEL KTPNNL EFL LPPHHGF AVKAS FRSEKHHNFGSRKFCETL KGVCVKASSSAL LEL VPETKKENIJ DFEL KTPNNL EFL LPPHHGF AVKAS FRSEKHHNFGSRK I CENWG KGVCVKASSSAL LEL VPETKKENIJ DFEL KTPNRL ELLYPLHELA KRHFL SPSPNPQNPNFKFFSRKPYQKKCRNGY I GVSSNQL L LOL VPETKKENIJ DFDL RTHNRL ELLYPLHELA KRHFL SPSPNPQNPNFKFFSRKPYQKKCRNGY I GVSSNQL L LOL VPETKKENIJ DFDL | |
| 40 | VIRFG HRYG HRYG SIFLA SIRYT NNLE NNLE NNLE NNSE | |
| | YSYRY LAYNI QQRR QQRR CQEKS CQECS CQECS CQETS CQETS CQETS CQETS CQETS CQETS CQETS CQETS CQETS CQETS CQETS CQETS CQETS CQETS CQUE | NVGL NVGL SVCS SVCS SVCS SVCS SVCS SVCS SVCS SVC |
| * | FPVVKRYSYRNIRFGL-CSVRASGGGSSGSSCSAVREDFADEEDFVKAGGSEILFVQMQQIKSMEVQAKLVDKIJPPIS NLSSSKLAYNIHRYGSSCRVDFQVRADGGSGSRTSVAYKEGFVDEEDFIKAGGSELLFVQMQQTKSMEKQAKLADKIPPIP NLSSSKLAYNIHRYGSSCRVDFQVRADGGSGSRTSVAYKEGFVDEEDFIKAGGSELLFVQMQQTKSMEKQAKLADKIPPIP HKFSLLKQRRFTNLSA-SSSLRQIKCSAKSDRCVVDKQGISVADEEDFIKAGGSELLFVQMQQTKSMEKQAKLADKIPPIP WSGGELCQEKSIFLAY-EQYESKCNSSSGSDSCVVDKEDFADEEDYIKAGGSELLFVQMQQTKSMESQSKLSEKIAQIP WSGGELCQEKSIFLAY-EQYESKCNSSSGSDSCVVDKEDFADEEDYIKAGGSELLFVQMQQTKSMESQSKLSDEIRQIS WSGGELCQEKSIFLAY-EQYESKCNSSSGSDSCVVDKEDFADEEDYIKAGGSELLFVQMQQTKSMESQSKLSDEIRQIS WSGGELCQEKSIFLAY-EQYESKCNSSSGSSSSSCVVDKEDFSSVVSGSSALLELVPETKKENIDFEL WDTLLRTHNKLEFLLPHHGFEKQHLVSTSKLQNQVFRIASRNIHPCRNGTVKARGSSALLELVPETKKENIDFEL WDTLLRTPNNLEFLHPVHGFSVKVSAFSSVKSQKFGAKKFCEGLGSRSVCVKASSSALLELVPETKKENIDFEL WDTLLKTPNKLEFLHPVHGFSVKASSFNSVKPHKFGSRKICENMGKGVCVKASSSALLELVPETKKENIDFEL WDTLLRTHNRLELLYPLHELAKRHFLSPSPNPQNPNFKFFSRKPYQKKCRNGYIGVSSNQLLDLVPETKKENIDFEL | 120 * 140 * DGALDARESAKLIĞI KYGL IĞDUĞ— ESWOLVVIĞCĞPAGLALAAESAKLIĞI KYGL IĞDUĞ— ESWOLVVIĞCĞPAGLSLAAEAAKLIĞI KYGL IĞDUĞ— ESWOLVVIĞCĞPAGLALAAESAKLIĞI KYGL IĞDUĞ— OTVLDL VVIĞCĞPAGLALAAESAKLIĞI KYGL IĞPDUĞ— OTVLDL VVIĞCĞPAGLALAAESAKLIĞI KYGL IĞPDUĞ— OTVLDL VVIĞCĞPAGLALAAESAKLIĞI KYGL IĞPDUĞ— SQVVÖLAVVĞCĞPAGLAVAQQVSEAĞL KYCS IDĞS—PKI SVVVÖLAVVĞCĞPAGLAVAQQVSEAĞL KYCS IDĞN—PKI SLVVÖLAVVĞCĞPAGLAVAQQVSEAĞL KYCS IDĞN—PKI NVVVÖLAVVĞCĞPAGLAVAQQVSEAĞL KYCS IDĞN—PKI NVVVÖLAVVĞCĞPAGLAVAQQVSEAĞL KYCS IDĞN—PKI NVVVÖLAVVĞCĞP SGLAVAQQVSEAĞL KYCS IDĞN—PKI |
| | STRNI | 140 AAESAI AAESAI AAESAI AAESAI AAESAI AAESAI AQQVSE AQQVSE AQQVSE AQQVSE AQQVSE |
| | \$ 5 5 5 6 5 6 5 6 5 6 5 6 5 6 5 6 5 6 5 | GLAV GLAL GLSL GLAL GLAV GLAV GLAV GLAV GLAV GLAV |
| 20 | SCPW SCPW TCPRI | * 8 8 8 8 8 8 8 8 8 8 |
| | MAVS - IS: WAVF | 120 DGALDHWYIGCGBAGLALAAE ESWDLWYIGCGBAGLSLAAE ESWDLWYIGCGBAGLSLAAE OTVLDLWYIGCGBAGLALAAE OTVLDLWYIGCGBAGLALAAE OTVLDLWYIGCGBAGLALAAE SQVVDLAVVGCGBAGLAVAQQ GVVVDLAVVGCGBAGLAVAQQ GVVVDLAVVGCGBAGLAVAQQ GVVVDLAVVGCGBAGLAVAQQ |
| * | MTATI | SOVODI DE LE SOVOD |
| | MECVGARNF-AAWAVSTFPSWS-CRRKFPVVKRYSYRNIRFGL-CSVRASGGGSSGSESCVAVREDFADEEDFVKAGGSE1LFVQXQQYKCMDEQSKLVDKIPPIS MELLGVRNLISSCPVWT-FGTRNLSSSKLAYNIHRYGSSCRVDPGVRAGGSGSSSSSVAYKEGFVDEEDFIKAGGSELLFVQXQQYKSMEKQAKLADKIPPIP MELLGVRNLISSCPVWT-FGTRNLSSSKLAYNIHRYGSSCRVDFQVRAGGSGSTSVAYKEGFVDEEDFIKAGGSELLFVQXQQYKSMEKQAKLADKIPPIP MECFGARWMTATWAVFTCPRFTDCNIRHKFSLLKQRRFTNLSA-SSSLRQIKCSAKSDRCVVDKQGISVADEEDFIKAGGSELFVQXQQKKOMQQQKKLSDEIRQIP MECVGARWMTATWAFTCPRFTDCNIRHKFSLLKQRRFTNLSA-SSSLRQIKCSAKSDRCVVDKQGISVADEEDFIKAGGSGLLFVQXQQKKOMQQQKKLSDEIRQIP MECVGARWMTATWAFTCPRFTM | 120 * 140 * 160 * 160 * 180 * 180 * 200 * |
| ΣĀ | | 120 * 140 * 160 * 160 * 180 * 200 * |
| 25 | | ** ** ** ** ** ** ** ** ** ** ** ** ** |
| (ה) ה | dopsi fel fel fel fel fopsi fopsi fel fel fel fel fel fel fel fel fel fel | E E E E E E E E E E E E E E E E E E E |
| FIG. 25A | ArabidopsisE AdonisEl AdonisE2 LettuceEE TomatoE ArabidopsisB AdonisB PepperB TomatoB TomatoB MarigoldB | PotatoE ArabidopsisE AdonisE1 AdonisE2 LettuceEE TomatoE ArabidopsisB AdonisB PepperB TomatoB PepperB TomatoB AdonisB DaffodilB |
| - à | | RA A A A A A A A A A A A A A A A A A A |

÷:

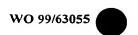
40/45 88 291 282 289 ĠŶŚYLSSKŶŒRITEAPNGYSLIEĞEGNITÎPCRLATVĄŚĠAAĞGKFLEŶELGGPRVCŸQTÄŶĞIEFĒŸENNŔŶŨPOĽŃVĠŊĠ---SKHKPESLEAKYPTFLŸŃ௸ ĨĠŶĨĹŶĹŊSKŶĎŖĨVEATNGQSĹVEĈŒĠŎŶVĨŢPĊŖFVŤVĀŠĠAĄŠĠĸFĹQŶĘĹĠSPŖŶSŲĠŢŔŶĠŶĔŶĎŊŇŖĠſŨĎŚĿŶŊŦŖŶŖĠŶ---ĹŔĦĎAQSĹEAKŶPŦFĹŶĄŴŖĄ EVDNNEEDPS LIVER FROM TROY --- VRHDAQS LEAKY PTFLY AMBA ĞŸSYLDSKÄERITEAGOGHSLVVÇENEIFÜPCRLATVÄSGAAĞGKLLEKEVGGPRVCVÇTAYĞVEVÜŞENNÜYÖPNÜÄVFMÜYRDİY——MQQKLQCSEEEYPTFLYÜMÜM GYSYLNSKYJERITEAGDGHSLVVÇENDIFİLPCRLATVASGAASGKLLEİZEVGGPRVCVQTAYGVEVEVEVENIPYÖPNLÜVFMOYROY ----MQQKLQCSEEEYPTFLYİMDM BYDPSUNVEMOYRDY --- TKHKSQSLEAQYPTFLYWNPM PYÖPDCMVEMDYRDY---TNEKVRSLEAEYPTFLYAMPM ſĠ-FŜRCLVQÑD-KPYNPGYQVÄYĞIVĄEVDGHPFQVDKŴFMDWRDKHLDSYPELKERNSKIPTFLYAMAF ĠŶĸFĦQAĸĸĨſĸvIH-EESKSLLĪĞNDGITĪJNATVVLDĀŢĠ-FŜRCLVQŶĎ-KPYNPGŸĞŶĀŸĠĪMĀĒŶĒEĦŖFĎĽDĶŶĹĠŊĸĠĸŔĠSHLNEKLELKDKNRKIPTFLŸĄŴŖF KKFHQAKKIKVIH-EESKSMLIČINGI TIOATVVLOATG-FÄRSLVQKO-KPYNPGYQVAYGI LAEVEEHREDVNKMVFMDARQSHLKNNVELKERNSRIPTFLYAMPF ŴĸFHQAKŴIKVIH-EESKSMLĽĈNDGI TĬŢQATVVLDÄŤG-FŜRSLVQĬD-KPYNPGYQVĄŶĞI LĄĔŸEEHŨFŨVNKÑVFNDÁRDSHLKNNTOLKERNSRĽPTFLYĄMPF ĠŶĸŦĦĦĦĸŶĨſĸvih-Eeaksmličndgvījgatvvldātg-FšrclvgŶD-kpykpgyġŷayĠtlaĒŸeehPfūtskávumomroshlgnnmelkernrkvptflyampf ĠŶĸFĦQAKŶĬſĸvih-Eelkslličndgvījgatlvldātg-FšrslvgŶD-kpynpgyġŶaŸĠtlaĔŸeeHPfōYDkŷJFmomroshldonletkarnsriptflyampf ĠĨ LAĔŸEEHĤFÛVDKŶLFMOWRDSHLDQNLEI KARNSR IPTFLYAMPF ĨĠĨĨŔŖĦŖŎŦĬŶĴŶĸŎĸĦ-EEEKSYLĬŹSDĠVŢĨĮĎARVŶĽĎſĬŖĬĠ-ŖŜŖĊĹŶĠŶĬŎ-KPŶŇPĠŶĨŶŶŶŖĨĹŹĘŶŒŦĤŔſĠĬŶŌŶŶŶŖŶŖŶŖŶŖŶŖŶŖŶŖ ĠŴSYLSSKŸĎS I TEASDGLRL VAČĎDNNVÍ PCRLATVÁSGAAŠGKLLQŶEVGGPRVCVÓTÁYĞVEVĚVENS <u>Ġ</u>ŴSYLSSK<u>Ŵ</u>ERITEAPNGLSLIEŒĠĠŊITĬPCRLATVĀŚĠĄĠĠĸLLQŶĘLGĠPRVCŲĢŤÁŶĠſĒVĒVĒSŢ GWLYLNSKNORIVEATNGHSLVECEGDVVIIPCRFVTVĀSGAĀGKFLONELGGPRVSVQTĀNGVEV OWKFHQSKÄTNVVH-EEANSTVVCSDGVKÄQASVVLDATG ArabidopsisE ArabidopsisB **AarigoldE** .ettuceEE **darigold**B)affodilB Adonis E2 Adonis E1 TomatoE AdonisB obaccoB PepperB [omatoB

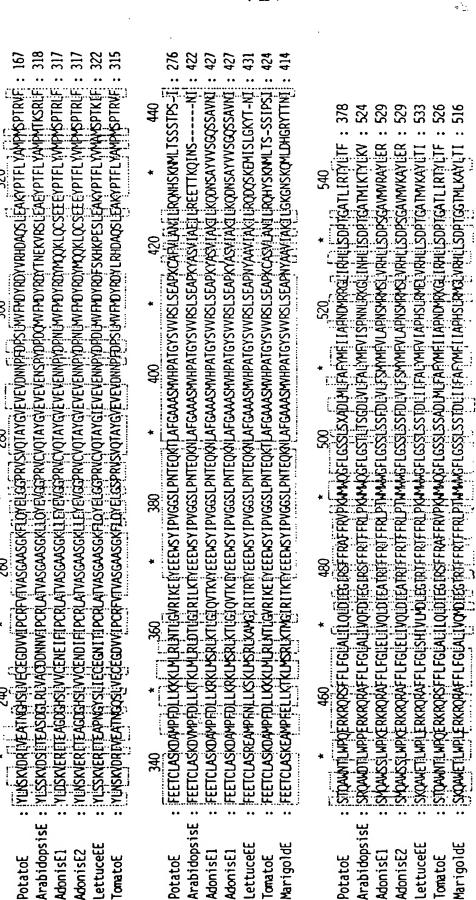
WO 99/63055

393 395 391 393 ENSYITHVEGSTRINTEGKTLAFGAAASAVHPATGYSVVRSLSEAPKCAFVLANTURONHSKNMLTSS TKSRĽFFEFICLASKOVMPFOLLKTKLMURUOTUGIRILKTYEEFISYIPYGGSLANTEOKNLAFGAAASMYHPATGYSYYRSUSEAPKYASVIAETUREETTKQINS--SPTRUFFEETCLASKDAMPFDLLKRKLMSRLIKTLIGI QVTKI VEEENSYI PIVGGSL PINTEQKNLAFGAAASNIVHPATGYSIVIRSI SEAPKYASVIAKILKQDNSAYVVSGQ SPTKI FFEET CLASREAMPFINLLKSKLMSRLIKAM GIRITRTY EFFMS YIT PLOG TOTAL FICA AND HIT FOR AND TOTAL STORY OF THE FEET CLASS OF THE STORY OF SPTRLFFEETCLASKDAMPFDLLKRKLMSRLIKTLIGIQVTKVYEFENSYI PVGGSLPNTEQKNLAFGAABNVHPATGYSVYRSLSEARKYABVTAK (LKQDNSAYVVSGQ SPTRYFFEETCLASKDAMPFDLLKKKLMURUNTUGVRIKEIYEEFUSYIHVGGSLANTEOKTLAFGAAASAVHPATGYSYVRSLSEAPKCASVLANTLROHYSKNMLTS-SPTKVIFIFETICI JASKEAMPFELLKTKLMSRLIKTMGI RITKTVIEFE JASYI PVISGSL PINTEOKNLAFIGA JASAN HPATGYSIVINSI SEAPINYAN IAKTLIGKGNSKOMLDHG SSNRTFILEETSLVARPGLRMEDIQERWAARIKHLGINVKRIEGDERCVI PYGGFILPVLPORVVGTGBTAGVVHPSTGYPVARTLAAAPIVANAVRYLGSPSSN----S SSTRÍFELE FÍSLVARPGLKMEDI QERMAVRLKHLÍGI KVKSTEEDERCVÍT FÍNGGÞL FÍVL ÞÓRVLGLÍGI FÁRMVHFSTGYMVÁRTLJAÁAFI VÁKSTI RYLNNEKSM---VAD KTLAAAPIVANSIVOYLVSDSGL----SSTKIFILETISLYARPGLRFEDIQERWARIKHLIGIKVKSIFFICKIFINGGPLPVLPQRVVGIGGTAGMVHPSTGMVARFILAAARVVAKSIVQYLGSDRSL----SSNKIFILEETSLYARPGLRMDDIQERMVÄRLINHLIGIKVKSIEFDEHOVIPMGGSLPVIPQRVVGTGGTÄGLVHPSTGYMVÄRTLÄAARVVÄNAIIHYLGSEKDL----SSNRIFILE FISLYARPGLOMDDIQE RAWARLSHIGIKVKSI EEDEHCVI PAGGPL PVL PQRVGIGG TAGAVHIST GYMVARTLAAA PVVANAI I QYLSSERSH----SSNRIFILEETSLYARPGLRIDDIQERAWARLAHLGIKVKSIEEDEHCLIFPKGGPLPIVLPQRVVGTGGTAGAYHPSTGYWARTLAAAAVVANAIIQYLGSERSH---SSNRIHEETSLVARPGLKMEDIQERNVARLINHIGIRIKSIEEDERCKIRKGEPLEVIRGERVIRKSIEEDER VIRAGERIAAN KOOTGESTAGENIYHE 340 FEEGGJASKDAMPFDLLKKKLMÜRÜNTÜĞVRIKETYÊ ArabidopsisE ArabidopsisB ettuceEE MariqoldE Daffodi 1B **Aarigold**B Adonis E2 PotatoE Idonis E1 FobaccoB AdonisB Pepper8 omatoE TómatoB

F16. 25B

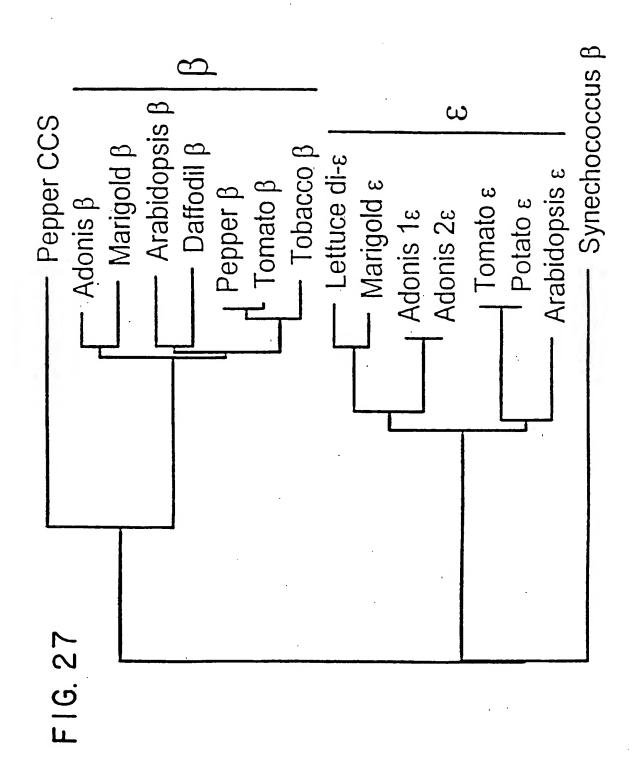
| | | مياجي | · - |
|---|--|---|---|
| 41/4 | 45 | | |
| 378 524 529 529 529 533 501 502 503 500 503 503 503 503 503 503 | : 103 : 102 : 102 : 107 : 100 : 84 | : 57 : 208 : 207 : 207 | : 212 : 205 : 194 |
| * "LTF | IPPIS IPPIP IPPIP IAQIP IRQIS | 220 VENEVI VESEVS CVESEVS | ESGVS ESGVS ESGVS |
| 0 TLIRTY TMIKTY TMIKTY TLIRTY TLIRTY TMINNI NMINNI | SKLVDK AKLADK AKLADK SKLSEK SKLSDE SSLSQK | | A X T |
| 540 SUPTGATI SUPTGATI SUPTGATI SUPTGATI TVP-LVNF | 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 | * 5 5 5 5 | |
| * LIRHLE LVRHCE | G C C C C C C C C C C C C C C C C C C C | GRVSR GRVSR GRVSR GRVSR | GRVHR GRVSR GRVSR |
| UDWRRGI UNLRKGI USWRYSI USWRKGI USREE UTSREE UTCRVEI UTCRVEI | SELFA SELFA SELFA SOLVA | 200 LIGRAY TIGRAY LIGRAY | RIGRAY LIGRAY LIGRAY |
| 520 MFV LAPP MFV LAPP MFV LAPP MFT I LAPP MF | VRASGGGSSGSESCVAVREDFADEEDFVKAGGSETTÉVOMONICIVADECISKÍ VDKLIPPTS VDFQVRADGGSGSRSSVAYKEGFVDEEDFTKAGGSETLÍFVQMONICIVADECISKÍ VDKLIPPTS VDFQVRADGGSGSRTSVAYKEGFVDEEDFTKAGGSETLÍFVQMONICIKSVEKQAKLADKLIPPTP SLRQTKCSAKSDRCVVDKQSTSVADEEDYVKAGGSETLFVQMONICIKSVESQSKLSEKLJAQTP YESKCNSSSGSDSCVVDKEDFADEEDYTKAGGSETLFVQMONICIVADAQSKLSDELRQTSKQTKCNAAKSQLVVKQETEEEEDYVKAGGSETLFVQMONICIVADAQSSTSSGSGVLIPRVP | * 200 * SOMEDE KOLGUQACTEHAMADITIVY LODODRI UTGRAVGRVSRH LHEEL URR GVWEDE FNOLGEQKCTEHVWRETTIVY LODOKRT TIGRAVGRVSRH LHEEL URR GVWEDE FKOLGERCTEHAMKOTTIVY LONDARV GRAVGRVSRH LHEEL URR | ODADPI ODOČEPI ODNOPI |
| MEALY VIFALY VIFSWY VIFSWY VIFSWY VIFEUSI VIFGUSI VIFGUSI VIFGUSI VIFGUSI | -ADEEC | | DTLVYL OTTVYL OTVVYL |
| SXADU SSFOL SSFOL SSFOL FLPEL FLPEL * | AVREDF- VYKEGF- VOKGGIS VOKEDF- VOKEDF- | 180 ACTEHVWF RCTEHVWF RCTEHAWK | I EHSK |
| MOGFIGSS MOGFIGSS MOGFIGSS MOGFIGSS MOGFIGSS MOGFIGSS MOGFIGSS MOGFIGSS MOGFIGSS MOGFIGSS MOGFIGSS MOGFISSS MOG | SSESCVA SSRSSVA SSRTSVA SSRTSVA SSDSCVA | רפוסאכ ורפוסאכ ורפובאכ | רפובפל |
| PKWY COC STANDS OF TWANTS | SGGGSSC SADGGSC SADGGSC CCSAKSE CCNSSSC CCNAAKS | DEFK COEFK COEFK | 60EFK |
| ARERU RIFFRL RIFFRL RIFFRL RAFFOL DAFFOL MFFOL MFFOL MFFOL MFFOL MFFOL MFFOL MFFOL MFFOL MFFOL MFFOL MFFOL MFFOL MFFOL | SVRAZ SVDEQVI SVDEQVI SSLRQII QYESI | 160 TNNYGVI TNNYGVI TNNYGVI | INNYGVI |
| GIRST GIRST GIRST GIRST GIRST GIRST GIRST GIRST GIRST GIRST GIRST GIRST GIRST GIRST GIRST | YRNIRFGL-C YNIHRYGSSC YNIHRYGSSC YNIHRYGSSC RRFTNLSA-S EKSIFLAY-E | SPOLPF | SPOLPF SPOLPF SPOLPF |
| STONAL STONAL STANAL ST | MECUGARNE-AAMAVSTFPSWS-CRRKFPVVKRYSYRNIRFGL-CSVRASGGGSSGSESCVAVREDFADEEDFYKAGGGELLFVQMQQIKSYEKGAKLADKLPPIS MELLGVRNLISSCPVMT-FGTRNLSSSKLAYNIHRYGSSCRVDFQVRADGGSGSRSSVAYKEGFVDEEDFIKAGGSELLFVQMQQIKSYEKQAKLADKLPPIP MELLGVRNLISSCPVWT-FGTRNLSSSKLAYNIHRYGSSCRVDFQVRADGGSGSRTSVAYKEGFVDEEDFIKAGGSELLFVQMQQIKSYEKQAKLADKLPPIP MECFGARNMTATWAVFTCPRFTDCNIRHKFSLLKQRRFTNLSA-SSSLRQIKCSAKSDRCVVDKQGISVADEEDYVKAGGSELFFVQMQRIKSYESGSKLSEKLAQIP MECVGVQNV-GAWAVLTRPRLNRWSGGELCQEKSIFLAY-EQYESKCNSSSGSDSCVVDKEDFADEEDYVKAGGSELLFVQMQQNKSYSYDAQSSLSDELRQIS MSMRAG-HMTATMAAFTCPRFMTSIRYTKQIKCNAAKSQLVVKQEIEEEEDYVKAGGSELLFVQMQQNKSYSDAQSSISQKLPRVP | * 200 * 220 GUKVGLIGPOLPFTNNYGVWEDEFNOLGLQACTEHVWRETTYYLODOKPITIGRAYGRVSRHLLHEELURRGVESGVS GUKVGLIGPOLPFTNNYGVWEDEFNOLGLGACTEHVWRETTYYLODOKPITIGRAYGRVSRRILHEELURRGVESGVS GUKVGLIGPOLPFTNNYGVWEDEFKOLGLERCTEHAWKOTTYYLONDAPVLIGRAYGRYSRHLLHEELUKRCVESGVS | LGENYGEIGPDLPFTNNYGVAQDEFIGLGGEGCIEHSWKDFLVYLODADPIRIGRAYGRYARDLLHEELLRRCYESGYS LGENYGLYGPDLPFTNNYGVWEDEFKDLGLQACIEHVWRDFIVYLODDEPILIGRAYGRYSRHFLHEELLKRCYENGYL LGENYAEIGPDLPFTNNYGVWEDEFIGLGLEGCIEHVWRDTVVYLODNDFILIGRAYGRYSRDLLHEELFTROYESGYS |
| FGALI | RKFPVKR STRNLSSSK STRNLSS STRN | O ESAKLG EAAKLG EAAKLG | ESAKLG ESAKLG ESAKLG |
| 2 | S-CRRK IT-FGTR IT-FGTR TDCNIR | 140 GLSLANE GLSLANE | 8 8 9 8 8 9 8 8 9 8 9 9 |
| 20 20 20 20 20 20 20 20 20 20 20 20 20 2 | STEPSING SSCPWN SSCPWN SSCPWN FTCPRE | 120 * 140 DGALOHVV IGCGPAGLALAAESAKI ESVADLVV IGCGPAGLSLAAEGAKI | NCTLGLVVTGCGPAGLALAAESAKI QTVLDLVVTGCGPAGLALAAESAKI JSNCTLGLVVTGCGPAGLALAGESAKI |
| * APPTILL APPT | | W WOLVY | |
| STPS-ISTQAWITLWPGERKRORSFFLFGLA SSAVNISRQAWDTLWPFERKRORAFFLFGLA SSAVNISRQAWDTLWPFERKRORAFFLFGLE SSAVNISRQAWFTLWPFERKRORAFFLFGLE KYT-NISRQAWFTLWPFERKRORAFFLFGLA RYTTNISRQAWFTLWPFERKRORAFFLFGLA RYTTNISRQAWFTLWPFERKRORAFFLFGLA SSIPSISTQAWNTLWPFERKRORAFFLFGLA SGDELSAAVWKDLWPFERRROREFFCFGWD -SGNELSTAVWKDLWPFERRROREFFCFGWD VTGDDLAAGIWRELWPFERRROREFFCFGWD -SGNDLSADVWKDLWPFERRROREFFCFGWD -SGNDLSADVWKDLWPFERRROREFFCFGWD -SGNDLSADVWKDLWPFERRROREFFCFGWD -SGNDLSADVWKDLWPFERRROREFFCFGWD | MECVGARNF-AAMAVSTFPSWS-CRRKFPVV MELLGVRNLISSCPVMT-FGTRNLSS MELLGVRNLISSCPVWT-FGTRNLSS MECFGARNMTATMAVFTCPRFTDCNIRHKFS MECVGVQNV-GAMAVLTRPRLNRWSGG | 120 06 ES | IGOVELLILIVVIGCGPAGLALAAESAKI NGOTVLDLVVIGCGPAGLALAAESAKI IGGGGDSNCTLIDLVVIGCGPAGLALAGESAKI |
| ω | | <u> </u> | ශි.ැද්ැයු: |
| F G. 2 Potatoc Arabidopsisc Adoniscil Lettucece Tomatoc Marigoldc Arabidopsis Pepper Tomatob Towatob Marigolds Marigolds PaffodilB PaffodilB | PotatoE ArabidopsisE AdonisEl AdonisE2 LettuceEE TomatoE MarigoldE | E opsisE E1 E2 | eeff E Jdf |
| F G. PotatoE Arabidops AdonisE1 LettuceEE TomatoE Arabidops AdonisB PepperB TomatoB TomatoB MarigoldB DaffodilB TomatoB TobaccoB | PotatoE Arabidops AdonisE1 AdonisE2 LettuceEE TomatoE | PotatoE ArabidopsisE AdonisE1 AdonisE2 | LettuceEE TomatoE MarigoldE |





F16. 26B

43/45



WO 99/63055





FIG. 28A

| Gap Weight: 12 Average Match: 2.912. Length Weight: 4 Average Mismatch: -2.003 Quality: 1837 Length: 534 Ratio: 3.499 Gaps: 3 Percent Similarity: 76.381 Percent Identity: 69.905 Match display thresholds for the alignment(s): = IDENTITY := 2 . = 1 |
|--|
| Arabidopsis x Lettuce |
| 1 MECVGARNF.AAMAVSTFPSWSCRRKFPVVKRYSYRNIRFGLCSVRA 46 46 |
| 47 SGGGSSGSESCVAVREDFADEEDFVKAGGSEILFVQMQQNKDMDEQSKLV 96 |
| 97 DKLPPISIGDGALDHVVIGCGPAGLALAAESAKLGLKVGLIGPDLPFTNN 146 : . |
| 147 YGVWEDEFNDLGLQKCIEHVWRETIVYLDDDKPITIGRAYGRVSRRLLHE 196 : :: : |
| 197 ELLRRCVESGVSYLSSKVDSİTEASDGLRLVACDDNNAIPCRLATVASGA 246 |
| 201 ÉLLRRCVÉSGVSÝLSSKVERÍTEÁPNGYSLIECEGNITÍPCRLÁTVÁSGÁ 250 |
| 247 ASGKLLQYEVGGPRVCVQTAYGVEVEVENSPYDPDQMVFMDYRDYTNEKV 296 : . : . |
| 297 RSLEAEYPTFLYAMPMTKSRLFFEETCLASKDVMPFDLLKTKLMLRLDTL 346 |
| 301 ESLEAKYPTFLYVMAMSPTKIFFEETCLASREAMPFNLLKSKLMSRLKAM 350 |

FIG. 28B

| 347 | GIRILKTYEEEWSYIPVGGSLPNTEQKNLAFGAAASMVHPATGYSVVRSL | 396 |
|-----|---|-----|
| 351 | : | 400 |
| 397 | SEAPKYASVIAEILREETTKQINSNISRQAWDTLWPPERKRQRAF | 441 |
| 401 | . . :: . :. : : SEAPNYAAVIAKILRQDQSKEMISLGKYTNISKQAWETLWPLERKRQRAF | 450 |
| 442 | FLFGLALIVOFDTEGIRSFFRTFFRLPKWMWQGFLGSTLTSGDLVLFALY | 491 |
| 451 | . . | 500 |
| | MFVISPNNLRKGLINHLISDPTGATMIKTYLKV* 525 | |
| 501 | . : : : : MFVIAPHSLRMELVRHLLSDPTGATMVKAYLTI* 534 | |



SEQUENCE LISTING

| <11 | .0> (| CUNN) SUN, | NGHA ZAIF | M JF REN | ļ., F | RANC | is x | ۲. | | | | | | | | |
|-----------------|------------------------------|--|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|-----|
| <12 | 0 < 0 ! | > GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND METHODS OF USE THEREOF | | | | | | | | | | | | | | |
| <13 | 0> 8 | 3172- | 9023 | | | | | | | | | | | | | |
| <14 <14 | 0> N 1> 1 | ЮТ Y .999- | ET A | SSIG | NED | | | | | | | | | | | |
| <15 <15 | 0> 0 1> 1 | 9/08 998- | 8,72 06-0 | 4 2 | | | | | | | | | | | | |
| <15 <15 | 0> 0 1> 1 | 9/08 998- | 8,72 06-0 | 5 2 | | | | | | | | | | | | |
| <16 | 0> 6 | 1 | | | | | | | | | | | | | | |
| <17 | 0> F | aten | tIn | Ver. | 2.0 | | | | | | | | | | | |
| <21 <21 | 0> 1 1> 1 2> D 3> A | 860 NA | dops | is t | hali. | ana | | | | | | | | | | |
| | 1> C | | (1 | 680) | | | | | | | | | | | | |
| | 0> 1 AAAG | GAA . | ATAA' | TTAG | AT TO | CCTC | TTTC' | T GC | TTGC | ТАТА | CCT | TGAT. | AGA . | ACAA | TATAAC | 60 |
| AAT(| GGTG | TAA | GTCT' | TCTC | GC T | GTAT' | rcga: | А АТ' | TATT' | TGGA | GGA | ggaa. | Me | | G TGT u Cys | 117 |
| STT /al | GGG Gly 5 | GCT Ala | AGG Arg | AAT Asn | TTC Phe | GCA Ala 10 | GCA Ala | ATG Met | GCG Ala | GTT Val | TCA Ser 15 | ACA Thr | TTT Phe | CCG Pro | TCA Ser | 165 |
| rgg rp 20 | AGT Ser | TGT Cys | CGA Arg | AGG Arg | AAA Lys 25 | TTT Phe | CCA Pro | GTG Val | GTT Val | AAG Lys 30 | AGA Arg | TAC Tyr | AGC Ser | TAT Tyr | AGG Arg 35 | 213 |
| AAT Asn | ATT Ile | CGT Arg | TTC Phe | GGT Gly 40 | TTG Leu | TGT Cys | AGT Ser | GTC Val | AGA Arg 45 | GCT Ala | AGC Ser | GGC Gly | GGC Gly | GGA Gly 50 | AGT Ser | 261 |
| CC Ser | GGT Gly | AGT Ser | GAG Glu 55 | AGT Ser | TGT Cys | GTA Val | GCG Ala | GTG Val 60 | AGA Arg | GAA Glu | GAT Asp | TTC Phe | GCT Ala 65 | GAC Asp | GAA Glu | 309 |
| AA Slu | GAT Asp | TTT Phe 70 | GTG Val | AAA Lys | GCT Ala | GGT Gly | GGT Gly 75 | TCT Ser | GAG Glu | ATT Ile | CTA Leu | TTT Phe 80 | GTT Val | CAA Gln | ATG Met | 357 |
| AG | CAG | AAC | AAA | GAT | ATG | GAT | GAA | CAG | тст | AAG | СТТ | GTT | САТ | AAG | ምም ር | 405 |

PCT/US99/12121

WO 99/63055



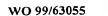
| Gln | Gln 85 | Asn | Lys | Asp | Met | Asp 90 | Glu | Gln | Ser | Lys | Leu 95 | Val | Asp | Lys | Leu | |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|------|
| CCT Pro 100 | CCT Pro | ATA Ile | TCA Ser | ATT Ile | GGT Gly 105 | GAT Asp | GGT Gly | GCT Ala | TTG Leu | GAT Asp 110 | His | GTG Val | GTT Val | ATT Ile | GGT Gly 115 | 453 |
| | | | | | Leu | | | | | | | | AAG Lys | | | 501 |
| TTA Leu | AAA Lys | GTT Val | GGA Gly 135 | CTC Leu | ATT Ile | GGT Gly | CCA Pro | GAT Asp 140 | CTT Leu | CCT Pro | TTT Phe | ACT Thr | AAC Asn 145 | AAT Asn | TAC Tyr | 549 |
| GGT Gly | GTT Val | TGG Trp 150 | GAA Glu | GAT Asp | GAA Glu | TTC Phe | AAT Asn 155 | GAT Asp | CTT Leu | GGG Gly | CTG Leu | CAA Gln 160 | AAA Lys | TGT Cys | ATT Ile | 597 |
| GAG Glu | CAT His 165 | GTT Val | TGG Trp | AGA Arg | GAG Glu | ACT Thr 170 | ATT Ile | GTG Val | TAT Tyr | CTG Leu | GAT Asp 175 | GAT Asp | GAC Asp | AAG Lys | CCT Pro | 645 |
| ATT Ile 180 | ACC Thr | ATT Ile | GGC Gly | CGT Arg | GCT Ala 185 | TAT Tyr | GGA Gly | AGA Arg | GTT Val | AGT Ser 190 | CGA Arg | CGT Arg | TTG Leu | CTC Leu | CAT His 195 | 693 |
| GAG Glu | GAG Glu | CTT Leu | TTG Leu | AGG Arg 200 | AGG Arg | TGT Cys | GTC Val | GAG Glu | TCA Ser 205 | GGT Gly | GTC Val | TCG Ser | TAC Tyr | CTT Leu 210 | AGC Ser | 741 |
| | | | | | | | | | | | | | AGA Arg 225 | | | 789 |
| GCT Ala | TGT Cys | GAC Asp 230 | GAC Asp | AAT Asn | AAC Asn | GTC Val | ATT Ile 235 | CCC Pro | TGC Cys | AGG Arg | CTT Leu | GCC Ala 240 | ACT Thr | GTT Val | GCT Ala | 837 |
| | | | | | | | | | | | | | GGT Gly | | | 885 |
| AGA Arg 260 | GTC Val | TGT Cys | GTG Val | CAA Gln | ACT Thr 265 | GCA Ala | TAC Tyr | GGC Gly | GTG Val | GAG Glu 270 | GTT Val | GAG Glu | GTG Val | GAA Glu | AAT Asn 275 | 933 |
| AGT Ser | CCA Pro | TAT Tyr | GAT Asp | CCA Pro 280 | GAT Asp | CAA Gln | ATG Met | GTT Val | TTC Phe 285 | ATG Met | GAT Asp | TAC Tyr | AGA Arg | GAT Asp 290 | TAT Tyr | 981 |
| | | | | | | | | | | | | | ACG Thr 305 | | | 1029 |
| | | | | | | | | | | | | | GAG Glu | | | 1077 |
| | | | | | | | | | | | | | ACG Thr | | | 1125 |

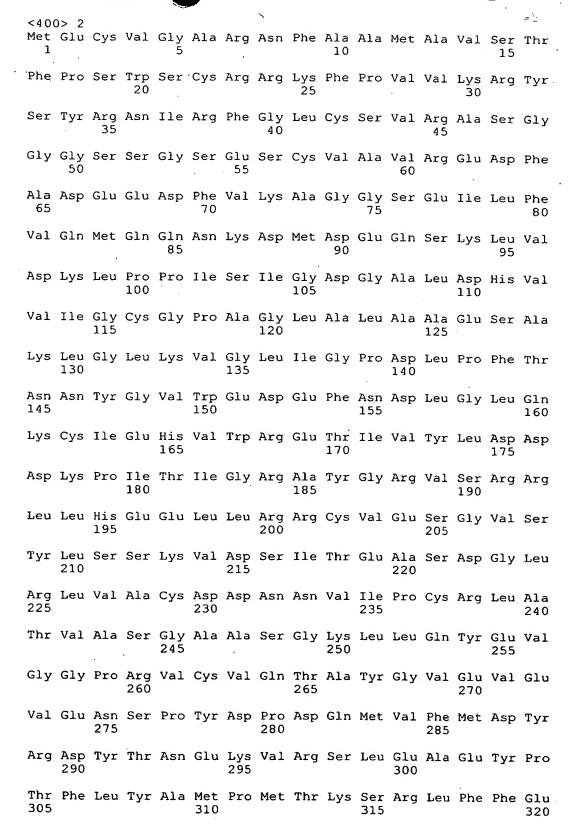
| WO | 99/63055 | • |
|----|----------|---|

PCT/US99/12121

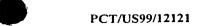
| 325 | | 330 | 33 | 35 | - 2 |
|---------------------------------------|-----------------------------------|-----------------------------------|---------------------------------|---|------------------------|
| ATG TTA AGA Met Leu Arg 340 | TTA GAT ACA Leu Asp Thr 345 | CTC GGA ATT Leu Gly Ile | CGA ATT CT Arg Ile Le 350 | TA AAG ACT TAC eu Lys Thr Tyr | GAA 1173 Glu 355 |
| GAG GAG TGG Glu Glu Trp | TCC TAT ATC Ser Tyr Ile 360 | CCA GTT GGT Pro Val Gly | GGT TCC TI Gly Ser Le 365 | IG CCA AAC ACC eu Pro Asn Thr 370 | GAA 1221 Glu |
| GIN Lys Asn | CTC GCC TTT Leu Ala Phe 375 | GGT GCT GCC Gly Ala Ala 380 | Ala Ser Me | TG GTA CAT CCC et Val His Pro 385 | GCA 1269 Ala |
| ACA GGC TAT Thr Gly Tyr 390 | TCA GTT GTG Ser Val Val | AGA TCT TTG Arg Ser Leu 395 | TCT GAA GC Ser Glu Al | CT CCA AAA TAT la Pro Lys Tyr 400 | GCA 1317 Ala |
| TCA GTC ATC Ser Val Ile 405 | GCA GAG ATA Ala Glu Ile | CTA AGA GAA Leu Arg Glu 410 | GAG ACT AC Glu Thr Th 41 | CC AAA CAG ATC or Lys Gln Ile 15 | AAC 1365 Asn |
| AGT AAT ATT Ser Asn Ile 420 | TCA AGA CAA Ser Arg Gln 425 | GCT TGG GAT Ala Trp Asp | ACT TTA TG Thr Leu Tr 430 | GG CCA CCA GAA | AGG 1413 Arg 435 |
| AAA AGA CAG . Lys Arg Gln . | AGA GCA TTC Arg Ala Phe 440 | TTT CTC TTT Phe Leu Phe | GGT CTT GC Gly Leu Al 445 | CA CTC ATA GTT La Leu Ile Val 450 | CAA 1461 Gln |
| Phe Asp Thr | GAA GGC ATT Glu Gly Ile 455 | AGA AGC TTC Arg Ser Phe 460 | TTC CGT AC | CT TTC TTC CGC or Phe Phe Arg 465 | CTT 1509 Leu |
| CCA AAA TGG 2 Pro Lys Trp 1 470 | ATG TGG CAA Met Trp Gln | GGG TTT CTA Gly Phe Leu 475 | GGA TCA AC Gly Ser Th | CA TTA ACA TCA or Leu Thr Ser 480 | GGA 1557 Gly |
| GAT CTC GTT (Asp Leu Val : 485 | CTC TTT GCT Leu Phe Ala | TTA TAC ATG Leu Tyr Met 490 | TTC GTC AT Phe Val II 49 | TT TCA CCA AAC e Ser Pro Asn | AAT 1605 Asn |
| Leu Arg Lys | GGT CTC ATC Gly Leu Ile 505 | Asn His Leu | ATC TCT GA Ile Ser As 510 | AT CCA ACC GGA | GCA 1653 Ala 515 |
| ACC ATG ATA A | AAA ACC TAT Lys Thr Tyr 520 | CTC AAA GTA Leu Lys Val | TGATTTACTT | ATCAACTCTT | 1700 |
| AGGTTTGTGT A | TATATATGT TO | GATTTATCT GA | ATAATCGA TC | CAAAGAATG GTATG | TGGGT 1760 |
| TACTAGGAAG T | TGGAAACAA AC | CATGTATAG AA | TCTAAGGA GT | GATCGAAA TGGAG | ATGGA 1820 |
| AACGAAAAGA A | AAAAATCAG TO | TTTGTTTT GT | GGTTAGTG | | 1860 |

<210> 2 <211> 524 <212> PRT <213> Arabidopsis thaliana





WO 99/63055



360

420

480

| Glu Thr Cys Leu Ala Ser Lys Asp Val Met Pro Phe Asp Leu Leu Lys 325 330 335 | | | | | | | | | | | |
|--|-----|--|--|--|--|--|--|--|--|--|--|
| Thr Lys Leu Met Leu Arg Leu Asp Thr Leu Gly Ile Arg Ile Leu Lys 340 345 350 | | | | | | | | | | | |
| Thr Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro 355 360 365 | | | | | | | | | | | |
| Asn Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val 370 375 380 | | | | | | | | | | | |
| His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro 385 390 395 400 | | | | | | | | | | | |
| Lys Tyr Ala Ser Val Ile Ala Glu Ile Leu Arg Glu Glu Thr Thr Lys 405 410 415 | | | | | | | | | | | |
| Gln Ile Asn Ser Asn Ile Ser Arg Gln Ala Trp Asp Thr Leu Trp Pro 420 425 430 | | | | | | | | | | | |
| Pro Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu Phe Gly Leu Ala Leu 435 440 445 | | | | | | | | | | | |
| Ile Val Gln Phe Asp Thr Glu Gly Ile Arg Ser Phe Phe Arg Thr Phe 450 455 460 | | | | | | | | | | | |
| Phe Arg Leu Pro Lys Trp Met Trp Gln Gly Phe Leu Gly Ser Thr Leu 465 470 475 480 | | | | | | | | | | | |
| Thr Ser Gly Asp Leu Val Leu Phe Ala Leu Tyr Met Phe Val Ile Ser 485 490 495 | | | | | | | | | | | |
| Pro Asn Asn Leu Arg Lys Gly Leu Ile Asn His Leu Ile Ser Asp Pro 500 505 510 | | | | | | | | | | | |
| Thr Gly Ala Thr Met Ile Lys Thr Tyr Leu Lys Val 515 | | | | | | | | | | | |
| <210> 3 <211> 956 <212> DNA <213> Arabidopsis thaliana | | | | | | | | | | | |
| <400> 3 GCTCTTTCTC CTCCTCCT ACCGATTTCC GACTCCGCCT CCCGAAATCC TTATCCGGAT | 60 | | | | | | | | | | |
| TCTCTCCGTC TCTTCGATTT AAACGCTTTT CTGTCTGTTA CGTCGTCGAA GAACGGAGAC | 120 | | | | | | | | | | |
| AGAATTCTCC GATTGAGAAC GATGAGAGAC CGGAGAGCAC GAGCTCCACA AACGCTATAG | 180 | | | | | | | | | | |
| ACGCTGAGTA TCTGGCGTTG CGTTTGGCGG AGAAATTGGA GAGGAAGAAA TCGGAGAGGT | 240 | | | | | | | | | | |
| CCACTTATCT AATCGCTGCT ATGTTGTCGA GCTTTTGGTAT CACTTCTATG GCTGTTATGG | 300 | | | | | | | | | | |

CTGTTTACTA CAGATTCTCT TGGCAAATGG AGGGAGGTGA GATCTCAATG TTGGAAATGT

TTGGTACATT TGCTCTCTC GTTGGTGCTG CTGTTGGTAT GGAATTCTGG GCAAGATGGG

CTCATAGAGC TCTGTGGCAC GCTTCTCTAT GGAATATGCA TGAGTCACAT CACAAACCAA

WO 99/63055



PCT/US99/12121

540

600

660

720

780

840 .

900 956

| | | | | | - | | | | | | | | | | |
|---|---|------------|------------|------------|------------|------------|------------|------------|------------|-------------------|------------|------------|------------|------------|------------|
| GAG | AAGG. | ACC | GTTT | GAGC' | TA A | ACGA' | rgtt: | T TT | GCTA' | TAGT | GAA | CGCT | GGT (| CCAG | CGATTG |
| GTC | TCCT | CTC | TATT | GGAT' | TC T' | rcaa' | TAAA | G GA | CTCG! | rtcc [°] | TGG | TCTC' | rgc ' | TTTG | GCGCC'G |
| GGT | TAGG | CAT . | AACG | GTGT' | TT G | GAAT(| CGCC' | r AC | ATGT: | ГТGТ | CCA | CGAT | GGT (| CTCG' | rgcaca |
| AGCGTTTCCC TGTAGGTCCC ATCGCCGACG TCCCTTACCT CCGAAAGGTC GCCGCCGCTC | | | | | | | | | | | | | | | |
| ACCAGCTACA TCACACAGAC AAGTTCAATG GTGTACCATA TGGACTGTTT CTTGGACCCA | | | | | | | | | | | | | | | |
| AGGAATTGGA AGAAGTTGGA GGAAATGAAG AGTTAGATAA GGAGATTAGT CGGAGAATCA | | | | | | | | | | | | | | | |
| AAT | CATA | CAA : | AAAG | GCCT | CG G | SCTC | CGGG: | r cg | AGTT | CGAG | TTC | rtga(| CTT | AAAT | CAAGTT |
| TTA | AATC | CCA . | AATT | CTTT' | TT T | rgtc | rtct | G TC | ATTA | rgat | CAT | CTTA | AGA (| CGGT | CT |
| <21 <21 | TTAAATCCCA AATTCTTTTT TTGTCTTCTG TCATTATGAT CATCTTAAGA CGGTCT <210> 4 <211> 294 <212> PRT <213> Arabidopsis thaliana | | | | | | | | | | | | | | |
| | 0> 4 | Ser | Sar | Sor | S-2- | Ωb ~ | 7 | Db - | 7 | 7 | . | | _ | _ | |
| 1 | 1110 | | 261 | 5 | Ser | 1111 | ASP | rne | 10 | ren | Arg | Leu | Pro | Lys 15 | Ser |
| Leu | Ser | Gly | Phe 20 | Ser | Pro | Ser | Leu | Arg 25 | Phe | Lys | Arg | Phe | Ser 30 | Val | Cys |
| Tyr | Val | Val 35 | Glu | Glu | Arg | Arg | Gln 40 | Asn | Ser | Pro | Ile | Glu 45 | Asn | Asp | Glu |
| Arg | Pro 50 | Glu | Ser | Thr | Ser | Ser 55 | Thr | Asn | Ala | Ile | Asp 60 | Ala | Glu | Tyr | Leu |
| Ala 65 | Leu | Arg | Leu | Ala | Glu 70 | Lys | Leu | Glu | Arg | Lys 75 | Lys | Ser | Glu | Arg | Ser 80 |
| Thr | Tyr | Leu | Ile | Ala 85 | Ala | Met | Leu | Ser | Ser 90 | Phe | Gly | Ile | Thr | Ser 95 | Met. |
| Ala | Val | Met | Ala 100 | Val | Tyr | Tyr | Arg | Phe 105 | Ser | Trp | Gln | Met | Glu 110 | Gly | Gly |
| Glu | Ile | Ser 115 | Met | Leu | Glu | Met | Phe 120 | Gly | Thr | Phe | Ala | Leu 125 | Ser | Val | Gly |
| Ala | Ala 130 | Val | Gly | Met | Glu | Phe 135 | Trp | Ala | Arg | Trp | Ala 140 | His | Arg | Ala | Leu |
| Trp 145 | His | Ala | Ser | Leu | Trp 150 | Met | Asn | His | Glu | Ser 155 | His | His | Lys | Pro | Arg 160 |
| Glu | Gly | Pro | Phe | Glu 165 | Leu | Asn | Asp | Val | Phe 170 | Ala | Ile | Val | Asn | Ala 175 | Gly |
| Pro | Ala | Ile | Gly 180 | Leu | Leu | Ser | Tyr | Gly 185 | Phe | Phe | Asn | Lys | Gly 190 | Leu | Val |

۲.

Pro Gly Leu Cys Phe Gly Ala Gly Leu Gly Ile Thr Val Phe Gly Ile 195 200



- Ala Tyr Met Phe Val His Asp Gly Leu Val His Lys Arg Phe Pro Val 210 220
- Gly Pro Ile Ala Asp Val Pro Tyr Leu Arg Lys Val Ala Ala Ala His 225 230 235 240
- Gln Leu His His Thr Asp Lys Phe Asn Gly Val Pro Tyr Gly Leu Phe 245 250 255
- Leu Gly Pro Lys Glu Leu Glu Glu Val Gly Gly Asn Glu Glu Leu Asp 260 265 270
- Lys Glu Ile Ser Arg Arg Ile Lys Ser Tyr Lys Lys Ala Ser Gly Ser 275 280 285
- Gly Ser Ser Ser Ser Ser 290
- <210> 5
- <211> 162
- <212> PRT
- <213> Alicalgenes sp.
- <400> 5
- Met Thr Gln Phe Leu Ile Val Val Ala Thr Val Leu Val Met Glu Leu 1 5 10 15
- Thr Ala Tyr Ser Val His Arg Trp Ile Met His Gly Pro Leu Gly Trp 20 25 30
- Gly Trp His Lys Ser His His Glu Glu His Asp His Ala Leu Glu Lys 35 40 45
- Asn Asp Leu Tyr Gly Val Val Phe Ala Val Leu Ala Thr Ile Leu Phe 50 55 60
- Thr Val Gly Ala Tyr Trp Trp Pro Val Leu Trp Trp Ile Ala Leu Gly 65 70 75 80
- Met Thr Val Tyr Gly Leu Ile Tyr Phe Ile Leu His Asp Gly Leu Val 85 90 95
- His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Arg Gly Tyr Phe Arg 100 105 110
- Arg Leu Tyr Gln Ala His Arg Leu His His Ala Val Glu Gly Arg Asp 115 120 125
- His Cys Val Ser Phe Gly Phe Ile Tyr Ala Pro Pro Val Asp Lys Leu 130 135 140
- Lys Gln Asp Leu Lys Arg Ser Gly Val Leu Arg Pro Gln Asp Glu Arg 145 150 155 160
- Pro Ser
- <210> 6
- <211> 175
- <212> PRT
- <213> Erwinia herbicola



Ile Ala Ala Phe Thr His Arg Tyr Ile Met His Gly Trp Gly Trp Arg
20 .25 30

Trp His Glu Ser His His Thr Pro Arg Lys Gly Val Phe Glu Leu Asn 35 40 45

Asp Leu Phe Ala Val Val Phe Ala Gly Val Ala Ile Ala Leu Ile Ala 50 55 60

Val Gly Thr Ala Gly Val Trp Pro Leu Gln Trp Ile Gly Cys Gly Met 65 70 75 80

Thr Val Tyr Gly Leu Leu Tyr Phe Leu Val His Asp Gly Leu Val His 85 90 95

Gln Arg Trp Pro Phe His Trp Ile Pro Arg Arg Gly Tyr Leu Lys Arg 100 105 110

Leu Tyr Val Ala His Arg Leu His His Ala Val Arg Gly Arg Glu Gly 115 120 125

Cys Val Ser Phe Gly Phe Ile Tyr Ala Arg Lys Pro Ala Asp Leu Gln 130 135 140

Ala Ile Leu Arg Glu Arg His Gly Arg Pro Pro Lys Arg Asp Ala Ala 145 150 155 160

Lys Asp Arg Pro Asp Ala Ala Ser Pro Ser Ser Ser Pro Glu 165 170 175

<210> 7

<211> 175

<212> PRT

<213> Erwinia uredovora

<400> 7

Met Leu Trp Ile Trp Asn Ala Leu Ile Val Phe Val Thr Val Ile Gly
1 5 10 15

Met Glu Val Ile Ala Ala Leu Ala His Lys Tyr Ile Met His Gly Trp 20 25 30

Gly Trp Gly Trp His Leu Ser His His Glu Pro Arg Lys Gly Ala Phe 35 40

Glu Val Asn Asp Leu Tyr Ala Val Val Phe Ala Ala Leu Ser Ile Leu 50 55 60

Leu Ile Tyr Leu Gly Ser Thr Gly Met Trp Pro Leu Gln Trp Ile Gly 65 70 75 80

Ala Gly Met Thr Ala Tyr Gly Leu Leu Tyr Phe Met Val His Asp Gly 85 90 95

Leu Val His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr 100 105 110

WO 99/63055



60

120

| Leu | Lys | Arg 115 | Leu | Tyr | Met | Ala | His 120 | Arg | Met | His | His | Ala 125 | Val | Arg | Gly |
|-------------|--------------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|---------------|------------|------------|------------|
| Lys | Glu 130 | Gly | Cys | Val | Ser | Phe 135 | Gly | Phe | Leu | Tyr | Ala 140 | Pro | Pro | Leu | Ser |
| Lys 145 | Leu | Gln | Ala | Thr | Leu 150 | Arg | Glu | Arg | His | Gly 155 | Ala | Arg | Ala | Gly | Ala 160 |
| Ala | Arg | Asp | Ala | Gln 165 | Gly | Gly | Glu | Asp | Glu 170 | Pro | Ala | Ser | Gly | Lys 175 | |
| <21: | 0> 8 1> 10 2> PI 3> A | RT | acte | rium | auri | ianti | icum | | | | | | | | , |
| | 0> 8 Thr | Asn | Phe | Leu 5 | Ile | Val | Val | Ala | Thr 10 | Val | Leu | Val | Met | Glu 15 | Leu |
| Thr | Ala | Tyr | Ser 20 | Val | His | Arg | Trp | Ile 25 | Met | His | Gly | Pro | Leu 30 | Gly | Trp |
| Gly | Trp | His 35 | Lys | Ser | His | His | Glu 40 | Glu | His | Asp | His | Ala 45 | Leu | Glu | Lys |
| Asn | Asp 50 | Leu | Tyr | Gly | Leu | Val 55 | Phe | Ala | Val | Ile | Ala 60 | Thr | Val | Leu | Phe |
| Thr 65 | Val | Gly | Trp | Ile | Trp 70 | Ala | Pro | Val | Leu | Trp 75 | Trp | Ile | Ala | Leu | Gly 80 |
| Met | Thr | Val | Tyr | Gly 85 | Leu | Ile | Tyr | Phe | Val 90 | Leu | His | Asp | Gly | Leu 95 | Val |
| His | Trp | Arg | Trp 100 | Pro | Phe | Arg | Tyr | Ile 105 | Pro | Arg | Lys | Gly | Tyr 110 | Ala | Arg |
| Arg | Leu | Tyr 115 | Gln | Ala | His | Arg | Leu 120 | His | His | Ala | Val | Glu 125 | Gly | Arg | Asp |
| His | Cys 130 | Val | Ser | Phe | Gly | | Ile | | | | Pro 140 | Val | Asp | Lys | Leu |
| Lys 145 | Gln | Asp | Leu | Lys | Met 150 | Ser | Gly | Val | Leu | Arg 155 | Ala | Glu | Ala | Gln | Glu 160 |
| Arg | Thr | | | | | | | | | | | | | | |
| <212 | l> 95 2> Di | AF | dopsi | is th | nalia | ana | | | | | | | | | |
| <400 CCA | _ | rcc (| SCCTO | cccc | ST TI | TTTT | rccga | A TCC | CGATO | CTCC | GGT | SCCG <i>P</i> | AGG P | ACTC | AGCTGT |

TTGTTCGCGC TTTCTCAGCC GTCACCATGA CCGATTCTAA CGATGCTGGA ATGGATGCTG

WO 99/63055 PCT/US99/12121

| TTCAGAGACG | ACTCATGTTT | GAAGACGAAT | GCATTCTCGT | TGATGAAAAT | AATCGTGTGG | 180 |
|--|-------------|------------|------------|------------|------------|-----|
| TGGGACATGA | CACTAAGTAT | AACTGTCATC | TGATGGAAAA | GATTGAAGCT | GAGAATTTAC | 240 |
| TTCACAGAGC | TTTCAGTGTG | TTTTTATTCA | ACTCCAAGTA | TGAGTTGCTT | CTCCAGCAAC | 300 |
| GGTCAAAAAC | AAAGGTTACT | TTCCCACTTG | TGTGGACAAA | CACTTGTTGC | AGCCATCCTC | 360 |
| TTTACCGTGA | ATCCGAGCTT | ATTGAAGAGA | ATGTGCTTGG | TGTAAGAAAT | GCCGCACAAA | 420 |
| GGAAGCTTTT | CGATGAGCTC | GGTATTGTAG | CAGAAGATGT | ACCAGTCGAT | GAGTTCACTC | 480 |
| CCTTGGGACG | CATGCTTTAC | AAGGCACCTT | CTGATGGGAA | ATGGGGAGAG | CACGAAGTTG | 540 |
| ACTATCTACT | CTTCATCGTG | CGGGATGTGA | AGCTTCAACC | AAACCCAGAT | GAAGTGGCTG | 600 |
| AGATCAAGTA | CGTGAGCAGG | GAAGAGCTTA | AGGAGCTGGT | GAAGAAAGCA | GATGCTGGCG | 660 |
| ATGAAGCTGT | GAAACTATCT | CCATGGTTCA | GATTGGTGGT | GGATAATTTC | TTGATGAAGT | 720 |
| GGTGGGATCA | TGTTGAGAAA | GGAACTATCA | CTGAAGCTGC | AGACATGAAA | ACCATTCACA | 780 |
| AGCTCTGAAC | TTTCCATAAG | TTTTGGATCT | TCCCCTTCCC | ATAATAAAT | TAAGAGATGA | 840 |
| GACTTTTATT | GATTACAGAC | AAAACTGGCA | ACAAAATCTA | TTCCTAGGAT | TTTTTTTTGC | 900 |
| ATTTATTTA | CTTTTGATTC | ATCTCTAGTT | TAGTTTTCAT | СТТАААААА | AAAA | 954 |
| <210> 10 <211> 996 <212> DNA <213> Arab | idopsis tha | liana | | · | | |
| <400> 10 CACCAATGTC | TGTTTCTTCT | ТТАТТТААТС | TCCCATTGAT | TCGCCTCAGA | TCTCTCGCTC | 60 |
| | | | | TCTGTCATCG | | 120 |
| | | | | GACAGATACT | | 180 |
| GTATGGATGC | TGTTCAGAGA | CGTCTCATGT | TTGAGGATGA | ATGCATTCTT | GTTGATGAAA | 240 |
| CTGATCGTGT | TGTGGGGCAT | GTCAGCAAGT | ATAATTGTCA | TCTGATGGAA | AATATTGAAG | 300 |
| CCAAGAATTT | GCTGCACAGG | GCTTTTAGTG | TATTTTTATT | CAACTCGAAG | TATGAGTTGC | 360 |
| TTCTCCAGCA | AAGGTCAAAC | ACAAAGGTTA | CGTTCCCTCT | AGTGTGGACT | AACACTTGTT | 420 |
| GCAGCCATCC | TCTTTACCGT | GAATCAGAGC | TTATCCAGGA | CAATGCACTA | GGTGTGAGGA | 480 |
| ATGCTGCACA | AAGAAAGCTT | CTCGATGAGC | TTGGTATTGT | AGCTGAAGAT | GTACCAGTCG | 540 |
| ATGAGTTCAC | TCCCTTGGGA | CGTATGCTGT | ACAAGGCTCC | TTCTGATGGC | AAATGGGGAG | 600 |
| | | | | | | |
| AGCATGAACT | TGATTACTTG | CTCTTCATCG | TGCGAGACGT | GAAGGTTCAA | CCAAACCCAG | 660 |

720

780

ATGAAGTAGC TGAGATCAAG TATGTGAGCC GGGAAGAGCT GAAGGAGCTG GTGAAGAAAG

CAGATGCAGG TGAGGAAGGT TTGAAACTGT CACCATGGTT CAGATTGGTG GTGGACAATT

BNSDOCID: <WO 996305541 IA>

| WO 99/63055 | |
|-------------|--|

| 4 | |
|---|--|
| 1 | |
| • | |

PCT/US99/12121

| TCTTGATGAA | GTGGTGGGAT | CATGTTGAGA | AAGGAACTTT | GGTTGAAGCT | ATAGACATGA | 840 |
|--|--------------|------------|------------|------------|------------|------|
| AAACCATCCA | CAAACTCTGA | ACATCTTTTT | TTAAAGTTTT | ТАААТСААТС | AACTTTCTCT | 900 |
| TCATCATTTT | TATCTTTTCG | ATGATAATAA | TTTGGGATAT | GTGAGACACT | TACAAAACTT | 960 |
| CCAAGCACCT | CAGGCAATAA | TAAAGTTTGC | GGCCGC | | | 996 |
| <210> 11 <211> 1165 <212> DNA <213> Haema | atococcus pi | luvialis | | | · | |
| | GGCCACAATC | GCTATTTGGA | ACCTGGCCCG | GCGGCAGTCC | GATGCCGCGA | 60 |
| TGCTTCGTTC | GTTGCTCAGA | GGCCTCACGC | ATATCCCCCG | CGTGAACTCC | GCCCAGCAGC | 120 |
| CCAGCTGTGC | ACACGCGCGA | CTCCAGTTTA | AGCTCAGGAG | CATGCAGATG | ACGCTCATGC | 180 |
| AGCCCAGCAT | CTCAGCCAAT | CTGTCGCGCG | CCGAGGACCG | CACAGACCAC | ATGAGGGGTG | 240 |
| CAAGCACCTG | GGCAGGCGGG | CAGTCGCAGG | ATGAGCTGAT | GCTGAAGGAC | GAGTGCATCT | 300 |
| TGGTGGATGT | TGAGGACAAC | ATCACAGGCC | ATGCCAGCAA | GCTGGAGTGT | CACAAGTTCC | 360 |
| TACCACATCA | GCCTGCAGGC | CTGCTGCACC | GGGCCTTCTC | TGTGTTCCTG | TTTGACGATC | 420 |
| AGGGGCGACT | GCTGCTGCAA | CAGCGTGCAC | GCTCAAAAAT | CACCTTCCCA | AGTGTGTGGA | 480 |
| CGAACACCTG | CTGCAGCCAC | CCTTTACATG | GGCAGACCCC | AGATGAGGTG | GACCAACTAA | 540 |
| GCCAGGTGGC | CGACGGAACA | GTACCTGGCG | CAAAGGCTGC | TGCCATCCGC | AAGTTGGAGC | 600 |
| ACGAGCTGGG | GATACCAGCG | CACCAGCTGC | CGGCAAGCGC | GTTTCGCTTC | CTCACGCGTT | 660 |
| TGCACTACTG | TGCCGCGGAC | GTGCAGCCAG | CTGCGACACA | ATCAGCGCTC | TGGGGCGAGC | 720 |
| ACGAAATGGA | CTACATCTTG | TTCATCCGGG | CCAACGTCAC | CTTGGCGCCC | AACCCTGACG | 780 |
| AGGTGGACGA | AGTCAGGTAC | GTGACGCAAG | AGGAGCTGCG | GCAGATGATG | CAGCCGGACA | 840 |
| ACGGGCTGCA | ATGGTCGCCG | TGGTTTCGCA | TCATCGCCGC | GCGCTTCCTT | GAGCGTTGGT | 900 |
| GGGCTGACCT | GGACGCGGCC | CTAAACACTG | ACAAACACGA | GGATTGGGGA | ACGGTGCATC | 960 |
| ACATCAACGA | AGCGTGAAAG | CAGAAGCTGC | AGGATGTGAA | GACACGTCAT | GGGGTGGAAT | 1020 |
| TGCGTACTTG | GCAGCTTCGT | ATCTCCTTTT | TCTGAGACTG | AACCTGCAGT | CAGGTCCCAC | 1080 |
| AAGGTCAGGT | AAAATGGCTC | GATAAAATGT | ACCGTCACTT | TTTGTCGCGT | ATACTGAACT | 1140 |
| CCAAGAGGTC | AAAAAAAAA | AAAAA | | | | 1165 |
| | | | | | | |

<210> 12 <211> 1135 <212> DNA

<213> Haematococcus pluvialis

PCT/US99/12121

| WO 99/63055 | |
|-------------|--|

| <400> 12 CTCGGTAGCT | GGCCACAATC | GCTATTTGGA | ACCTGGCCCG | GCGGCAGTCC | GATGCCGCGA | 60 |
|---|------------|------------|--|------------|------------|------|
| TGCTTCGTTC | GTTGCTCAGA | GGCCTCACGC | ATATCCCGCG | CGTGAACTCC | GCCCAGCAGC | 120 |
| CCAGCTGTGC | ACACGCGCGA | CTCCAGTTTA | AGCTCAGGAG | CATGCAGCTG | CTTTCCGAGG | 180 |
| ACCGCACAGA | CCACATGAGG | GGTGCAAGCA | CCTGGGCAGG | CGGGCAGTCG | CAGGATGAGC | 240 |
| TGATGCTGAA | GGACGAGTGC | ATCTTGGTAG | ATGTTGAGGA | CAACATCACA | GGCCATGCCA | 300 |
| GCAAGCTGGA | GTGTCACAAG | TTCCTACCAC | ATCAGCCTGC | AGGCCTGCTG | CACCGGGCCT | 360 |
| TCTCTGTGTT | CCTGTTTGAC | GATCAGGGGC | GACTGCTGCT | GCAACAGCGT | GCACGCTCAA | 420 |
| AAATCACCTT | CCCAAGTGTG | TGGACGAACA | CCTGCTGCAG | CCACCCTTTA | CATGGGCAGA | 480 |
| CCCCAGATGA | GGTGGACCAA | CTAAGCCAGG | TGGCCGACGG | AACAGTACCT | GGCGCAAAGG | 540 |
| CTGCTGCCAT | CCGCAAGTTG | GAGCACGAGC | TGGGGATACC | AGCGCACCAG | CTGCCGGCAA | 600 |
| GCGCGTTTCG | CTTCCTCACG | CGTTTGCACT | ACTGTGCCGC | GGACGTGCAG | CCAGCTGCGA | 660 |
| CACAATCAGC | GCTCTGGGGC | GAGCACGAAA | TGGACTACAT | CTTGTTCATC | CGGGCCAACG | 720 |
| TCACCTTGGC | GCCCAACCCT | GACGAGGTGG | ACGAAGTCAG | GTÄCGTGACG | CAAGAGGAGC | 780 |
| TGCGGCAGAT | GATGCAGCCG | GACAACGGGC | TTCAATGGTC | GCCGTGGTTT | CGCATCATCG | 840 |
| CCGCGCGCTT | CCTTGAGCGT | TGGTGGGCTG | ACCTGGACGC | GGCCCTAAAC | ACTGACAAAC | 900 |
| ACGAGGATTG | GGGAACGGTG | CATCACATCA | ACGAAGCGTG | AAGGCAGAAG | CTGCAGGATG | 960 |
| TGAAGACACG | TCATGGGGTG | GAATTGCGTA | CTTGGCAGCT | TCGTATCTCC | TTTTTCTGAG | 1020 |
| ACTGAACCTG | CAGAGCTAGA | GTCAATGGTG | CATCATATTC | ATCGTCTCTC | TTTTGTTTTA | 1080 |
| GACTAATCTG | TAGCTAGAGT | CACTGATGAA | TCCTTTACAA | CTTTCAAAAA | AAAAA | 1135 |
| <210> 13 <211> 960 <212> DNA <213> Taget | ces erecta | | | | | |
| <400> 13 | СТСАААТСТС | CTCCCTCCCT | · • • • • • • • • • • • • • • • • • • • | CATCCCTCAC | Ch Cmaaaaa | 60 |
| | TCAGCGACGT | 4 | | • | | 60 |
| | GGGACATGAT | | | | | 120 |
| | GCACAGAGCA | | | | | 180 |
| | | | | | | 240 |
| | GTCTGCAACC | | | | | 300 |
| * | CTACAGAGAA | | | | | 360 |
| | NNNNNNNNN | | • | | | 420 |
| иииииииии | иииииииии | иииииииии | иииииииии | иииииииии | иииииииии | 480 |

| , |
|---|
|---|

| WO 99/63055 | | | , | PCT/US99/12121 |
|-----------------------------------|-----------|-----------|-----------|----------------|
| и инининини инининини инининини | иииииииии | инининини | ทททททททท | N 540 |
| и пинининини пинининини инининини | имимимими | иииииииии | иииииииии | N 600 |

| | | | | | _ | |
|------------|------------|------------|------------|------------|------------|-----|
| ииииииииии | ииииииииии | ииииииииии | ииииииииии | иииииииии | имимимими | 600 |
| ииииииииии | иииииииии | имимимими | имимимими | иииииииии | иииииииии | 660 |
| имимимими | ииииииииии | TCATGTGCAA | AAGGGTACAC | TCACTGAATG | CAATTTGATA | 720 |
| TGAAAACCAT | ACACAAGCTG | ATATAGAAAC | ACACCCTCAA | CCGAAAAGCA | AGCCTAATAA | 780 |
| TTCGGGTTGG | GTCGGGTCTA | CCATCAATTG | TTTTTTTTTT | TTAACAACTT | TTAATCTCTA | 840 |
| TTTGAGCATG | TTGATTCTTG | TCTTTTGTGT | GTAAGATTTT | GGGTTTCGTT | TCAGTTGTAA | 900 |
| TAATGAACCA | TTGATGGTTT | GCAATTTCAA | GTTCCTATCG | ACATGTAGTG | ממממממדידמ | 960 |

<210> 14

<211> 305

<212> PRT

<213> Haematococcus pluvialis

<400> 14

Met Leu Arg Ser Leu Leu Arg Gly Leu Thr His Ile Pro Arg Val Asn

Ser Ala Gln Gln Pro Ser Cys Ala His Ala Arg Leu Gln Phe Lys Leu

Arg Ser Met Gln Met Thr Leu Met Gln Pro Ser Ile Ser Ala Asn Leu

Ser Arg Ala Glu Asp Arg Thr Asp His Met Arg Gly Ala Ser Thr Trp

Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys Asp Glu Cys Ile

Leu Val Asp Val Glu Asp Asn Ile Thr Gly His Ala Ser Lys Leu Glu

Cys His Lys Phe Leu Pro His Gln Pro Ala Gly Leu Leu His Arg Ala 105

Phe Ser Val Phe Leu Phe Asp Asp Gln Gly Arg Leu Leu Gln Gln

Arg Ala Arg Ser Lys Ile Thr Phe Pro Ser Val Trp Thr Asn Thr Cys 135

Cys Ser His Pro Leu His Gly Gln Thr Pro Asp Glu Val Asp Gln Leu

Ser Gln Val Ala Asp Gly Thr Val Pro Gly Ala Lys Ala Ala Ile

Arg Lys Leu Glu His Glu Leu Gly Ile Pro Ala His Gln Leu Pro Ala 180 185

Ser Ala Phe Arg Phe Leu Thr Arg Leu His Tyr Cys Ala Ala Asp Val 200



Gln Pro Ala Ala Thr Gln Ser Ala Leu Trp Gly Glu His Glu Met Asp 210 215 220

Tyr Ile Leu Phe Ile Arg Ala Asn Val Thr Leu Ala Pro Asn Pro Asp 225 230 235 240

Glu Val Asp Glu Val Arg Tyr Val Thr Gln Glu Glu Leu Arg Gln Met 245 250 255

Met Gln Pro Asp Asn Gly Leu Gln Trp Ser Pro Trp Phe Arg Ile Ile 260 265 270

Ala Ala Arg Phe Leu Glu Arg Trp Trp Ala Asp Leu Asp Ala Ala Leu 275 280 285

Asn Thr Asp Lys His Glu Asp Trp Gly Thr Val His His Ile Asn Glu 290 295 300

Ala 305

<210> 15

<211> 293

<212> PRT

<213> Haematococcus pluvialis

<400> 15

Met Leu Arg Ser Leu Leu Arg Gly Leu Thr His Ile Pro Arg Val Asn
1 5 10 15

Ser Ala Gln Gln Pro Ser Cys Ala His Ala Arg Leu Gln Phe Lys Leu 20 25 30

Arg Ser Met Gln Leu Leu Ser Glu Asp Arg Thr Asp His Met Arg Gly 35 40 45

Ala Ser Thr Trp Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys
50 55 60

Asp Glu Cys Ile Leu Val Asp Val Glu Asp Asn Ile Thr Gly His Ala 65 70 75 80

Ser Lys Leu Glu Cys His Lys Phe Leu Pro His Gln Pro Ala Gly Leu 85 90 95

Leu His Arg Ala Phe Ser Val Phe Leu Phe Asp Asp Gln Gly Arg Leu 100 105 110

Leu Leu Gln Gln Arg Ala Arg Ser Lys Ile Thr Phe Pro Ser Val Trp
115 120 125

Thr Asn Thr Cys Cys Ser His Pro Leu His Gly Gln Thr Pro Asp Glu 130 135 140

Val Asp Gln Leu Ser Gln Val Ala Asp Gly Thr Val Pro Gly Ala Lys
145 150 155 160

Ala Ala Ile Arg Lys Leu Glu His Glu Leu Gly Ile Pro Ala His 165 170 175

Gln Leu Pro Ala Ser Ala Phe Arg Phe Leu Thr Arg Leu His Tyr Cys

180 185 190

Ala Ala Asp Val Gln Pro Ala Ala Thr Gln.Ser Ala Leu Trp Gly Glu 195 200 205

His Glu Met Asp Tyr Ile Leu Phe Ile Arg Ala Asn Val Thr Leu Ala 210 215 220

Pro Asn Pro Asp Glu Val Asp Glu Val Arg Tyr Val Thr Gln Glu Glu 225 230 235 240

Leu Arg Gln Met Met Gln Pro Asp Asn Gly Leu Gln Trp Ser Pro Trp 245 250 255

Phe Arg Ile Ile Ala Ala Arg Phe Leu Glu Arg Trp Trp Ala Asp Leu 260 265 270

Asp Ala Ala Leu Asn Thr Asp Lys His Glu Asp Trp Gly Thr Val His 275 280 285

His Ile Asn Glu Ala 290

<210> 16

<211> 284

<212> PRT

<213> Arabidopsis thaliana

<400> 16

Met Ser Val Ser Ser Leu Phe Asn Leu Pro Leu Ile Arg Leu Arg Ser 1 5 10 15

Leu Ala Leu Ser Ser Ser Phe Ser Ser Phe Arg Phe Ala His Arg Pro
20 25 30

Leu Ser Ser Ile Ser Pro Arg Lys Leu Pro Asn Phe Arg Ala Phe Ser 35 40 45

Gly Thr Ala Met Thr Asp Thr Lys Asp Ala Gly Met Asp Ala Val Gln 50 60

Arg Arg Leu Met Phe Glu Asp Glu Cys Ile Leu Val Asp Glu Thr Asp 65 70 75 80

Arg Val Val Gly His Val Ser Lys Tyr Asn Cys His Leu Met Glu Asn 85 90 95

Ile Glu Ala Lys Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe 100 105 110

Asn Ser Lys Tyr Glu Leu Leu Gln Gln Arg Ser Asn Thr Lys Val 115 120 125

Thr Phe Pro Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr 130 135 140

Arg Glu Ser Glu Leu Ile Gln Asp Asn Ala Leu Gly Val Arg Asn Ala 145 150 . 155 160

Ala Gln Arg Lys Leu Leu Asp Glu Leu Gly Ile Val Ala Glu Asp Val 165 170 175



Pro Val Asp Glu Phe Thr Pro Leu Gly Arg Met Leu Tyr Lys Ala Pro 180 . 185 190

Ser Asp Gly Lys Trp Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Ile 195 200 205

Val Arg Asp Val Lys Val Gln Pro Asn Pro Asp Glu Val Ala Glu Ile 210 215 220

Lys Tyr Val Ser Arg Glu Glu Leu Lys Glu Leu Val Lys Lys Ala Asp 225 230 235 240

Ala Gly Glu Glu Leu Lys Leu Ser Pro Trp Phe Arg Leu Val Val 245 250 255

Asp Asn Phe Leu Met Lys Trp Trp Asp His Val Glu Lys Gly Thr Leu 260 265 270

Val Glu Ala Ile Asp Met Lys Thr Ile His Lys Leu 275 280

<210> 17

<211> 287

<212> PRT

<213> Clarkia breweri

<400> 17

Met Ser Ser Ser Met Leu Asn Phe Thr Ala Ser Arg Ile Val Ser Leu
1 5 10 15

Pro Leu Leu Ser Ser Pro Pro Ser Arg Val His Leu Pro Leu Cys Phe 20 25 30

Phe Ser Pro Ile Ser Leu Thr Gln Arg Phe Ser Ala Lys Leu Thr Phe 35 40 45

Ser Ser Gln Ala Thr Thr Met Gly Glu Val Val Asp Ala Gly Met Asp 50 60

Ala Val Gln Arg Arg Leu Met Phe Glu Asp Glu Cys Ile Leu Val Asp 65 70 75 80

Glu Asn Asp Lys Val Val Gly His Glu Ser Lys Tyr Asn Cys His Leu 85 90 95

Met Glu Lys Ile Glu Ser Glu Asn Leu Leu His Arg Ala Phe Ser Val 100 105 110

Phe Leu Phe Asn Ser Lys Tyr Glu Leu Leu Gln Gln Arg Ser Ala 115 120 125

Thr Lys Val Thr Phe Pro Leu Val Trp Thr Asn Thr Cys Cys Ser His 130 135 140

Pro Leu Tyr Arg Glu Ser Glu Leu Ile Asp Glu Asn Cys Leu Gly Val 145 150 155 160

Arg Asn Ala Ala Gln Arg Lys Leu Leu Asp Glu Leu Gly Ile Pro Ala 165 170 175

Glu Asp Leu Pro Val Asp Gln Phe Ile Pro Leu Ser Arg Ile Leu Tyr

180 185

Lys Ala Pro Ser Asp Gly Lys Trp Gly Glu His Glu Leu Asp Tyr Leu 195 200 205

Leu Phe Ile Ile Arg Asp Val Asn Leu Asp Pro Asn Pro Asp Glu Val 210 215 220

Ala Glu Val Lys Tyr Met Asn Arg Asp Asp Leu Lys Glu Leu Leu Arg 225 230 235 240

Lys Ala Asp Ala Glu Glu Glu Gly Val Lys Leu Ser Pro Trp Phe Arg 245 250 255

Leu Val Val Asp Asn Phe Leu Phe Lys Trp Trp Asp His Val Glu Lys 260 265 270

Gly Ser Leu Lys Asp Ala Ala Asp Met Lys Thr Ile His Lys Leu 275 280 285

<210> 18

<211> 261

<212> PRT

<213> Arabidopsis thaliana

<400> 18

Thr Gly Pro Pro Pro Arg Phe Phe Pro Ile Arg Ser Pro Val Pro Arg
1 5 10 15

Thr Gln Leu Phe Val Arg Ala Phe Ser Ala Val Thr Met Thr Asp Ser 20 25 30

Asn Asp Ala Gly Met Asp Ala Val Gln Arg Arg Leu Met Phe Glu Asp $35 \hspace{1cm} 40 \hspace{1cm} 45$

Glu Cys Ile Leu Val Asp Glu Asn Asn Arg Val Val Gly His Asp Thr 50 55 60

Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala Glu Asn Leu Leu 65 70 75 80

His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu Leu Leu 85 90 95

Leu Gln Gln Arg Ser Lys Thr Lys Val Thr Phe Pro Leu Val Trp Thr
100 105 110

Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser Glu Leu Ile Glu 115 120 125

Glu Asn Val Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu Phe Asp 130 135 140

Glu Leu Gly Ile Val Ala Glu Asp Val Pro Val Asp Glu Phe Thr Pro 145 150 155 160

Leu Gly Arg Met Leu Tyr Lys Ala Pro Ser Asp Gly Lys Trp Gly Glu
165 170 175

His Glu Val Asp Tyr Leu Leu Phe Ile Val Arg Asp Val Lys Leu Gln 180 185 190





Pro Asn Pro Asp Glu Val Ala Glu Ile Lys Tyr Val Ser Arg Glu Glu 195 200 205

Leu Lys Glu Leu Val Lys Lys Ala Asp Ala Gly Asp Glu Ala Val Lys 210 220

Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe Leu Met Lys Trp 225 230 235 240

Trp Asp His Val Glu Lys Gly Thr Ile Thr Glu Ala Ala Asp Met Lys 245 250 255

Thr Ile His Lys Leu 260

<210> 19

<211> 288

<212> PRT

<213> Saccharomyces cerevisiae

<400> 19

Met Thr Ala Asp Asn Asn Ser Met Pro His Gly Ala Val Ser Ser Tyr 1 5 10 15

Ala Lys Leu Val Gl
n Asn Gl
n Thr Pro Glu Asp Ile Leu Glu Glu Phe 20
 25
 30

Pro Glu Ile Ile Pro Leu Gln Gln Arg Pro Asn Thr Arg Ser Ser Glu 35 40 45

Thr Ser Asn Asp Glu Ser Gly Glu Thr Cys Phe Ser Gly His Asp Glu 50 55 60

Glu Gln Ile Lys Leu Met Asn Glu Asn Cys Ile Val Leu Asp Trp Asp 65 70 75 80

Asp Asn Ala Ile Gly Ala Gly Thr Lys Lys Val Cys His Leu Met Glu 85 90 95

Asn Ile Glu Lys Gly Leu Leu His Arg Ala Phe Ser Val Phe Ile Phe 100 105 110

Asn Glu Gln Gly Glu Leu Leu Gln Gln Arg Ala Thr Glu Lys Ile 115 120 125

Thr Phe Pro Asp Leu Trp Thr Asn Thr Cys Cys Ser His Pro Leu Cys 130 135 140

Ile Asp Asp Glu Leu Gly Leu Lys Gly Lys Leu Asp Asp Lys Ile Lys 145 150 155 160

Gly Ala Ile Thr Ala Ala Val Arg Lys Leu Asp His Glu Leu Gly Ile 165 170 175

Pro Glu Asp Glu Thr Lys Thr Arg Gly Lys Phe His Phe Leu Asn Arg 180 185 190

Ile His Tyr Met Ala Pro Ser Asn Glu Pro Trp Gly Glu His Glu Ile 195 200 205

Asp Tyr Ile Leu Phe Tyr Lys Ile Asn Ala Lys Glu Asn Leu Thr Val

Asn Pro Asn Val Asn Glu Val Arg Asp Phe Lys Trp Val Ser Pro Asn 225 230 235 240

220

Asp Leu Lys Thr Met Phe Ala Asp Pro Ser Tyr Lys Phe Thr Pro Trp 245 250 255

Phe Lys Ile Ile Cys Glu Asn Tyr Leu Phe Asn Trp Trp Glu Gln Leu 260 265 270

Asp Asp Leu Ser Glu Val Glu Asn Asp Arg Gln Ile His Arg Met Leu 275 280 285

<210> 20

<211> 456

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Consensus
 sequence of four plant B-cyclases

<400> 20

Met Asp Thr Leu Leu Lys Thr Pro Asn Leu Glu Phe Leu Pro His Gly
1 5 10 15

Phe Val Lys Ser Phe Ser Lys Phe Gly Lys Cys Glu Gly Val Cys Val 20 25 30

Lys Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr Lys Lys Glu Asn 35 40 45

Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys Gly Val Val Asp
50 55 60

Leu Ala Val Val Gly Gly Pro Ala Gly Leu Ala Val Ala Gln Gln 65 70 75 80'

Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile Asp Pro Pro Lys Leu 85 90 95

Ile Trp Pro Asn Asn Tyr Gly Val Trp Val Asp Glu Phe Glu Ala Met 100 105 110

Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp Ser Gly Ala Val Tyr Ile 115 120 125

Asp Asp Thr Lys Asp Leu Arg Pro Tyr Gly Arg Val Asn Arg Lys Gln 130 135 140

Leu Lys Ser Lys Met Met Gln Lys Cys Ile Asn Gly Val Lys Phe His 145 150 155 160

Gln Ala Lys Val Ile Lys Val Ile His Glu Glu Lys Ser Met Leu Ile 165 170 175

Cys Asn Asp Gly Thr Ile Gln Ala Thr Val Val Leu Asp Ala Thr Gly 180 185 190

Phe Ser Arg Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln

A ...

195 200 205

Val Ala Tyr Gly Ile Leu Ala Glu Val Glu Glu His Pro Phe Asp Lys 210 220

Met Val Phe Met Asp Trp Arg Asp Ser His Leu Asn Asn Glu Leu Lys 235 240

Glu Arg Asn Ser Ile Pro Thr Phe Leu Tyr Ala Met Pro Phe Ser Ser 245 250 255

Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg Pro Gly Leu 260 265 270

Arg Met Asp Asp Ile Gln Glu Arg Met Val Ala Arg Leu His Leu Gly 275 280 285

Ile Lys Val Lys Ser Ile Glu Glu Asp Glu His Cys Val Ile Pro Met 290 295 300

Gly Gly Pro Leu Pro Val Leu Pro Gln Arg Val Val Gly Ile Gly Gly 305 310 315 320

Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met Val Ala Arg Thr 325 330 335

Leu Ala Ala Pro Val Val Ala Asn Ala Ile Ile Tyr Leu Gly Ser 340 345 350

Glu Ser Ser Gly Glu Leu Ser Ala Glu Val Trp Lys Asp Leu Trp Pro 355 360 365

Ile Glu Arg Arg Arg Gln Arg Glu Phe Phe Cys Phe Gly Met Asp Ile 370 375 380

Leu Leu Lys Leu Asp Leu Pro Ala Thr Arg Arg Phe Phe Asp Ala Phe 385 390 395 400

Phe Asp Leu Glu Pro Arg Tyr Trp His Gly Phe Leu Ser Ser Arg Leu 405 410 415

Phe Leu Pro Glu Leu Ile Val Phe Gly Leu Ser Leu Phe Ser His Ala 420 425 430

Ser Asn Thr Ser Arg Glu Ile Met Thr Lys Gly Thr Pro Leu Val Met 435 440 445

Ile Asn Asn Leu Leu Gln Asp Glu 450 455

<210> 21

<211> 524

<212> PRT

<213> Arabidopsis thaliana

<400> 21

Met Glu Cys Val Gly Ala Arg Asn Phe Ala Ala Met Ala Val Ser Thr 1 5 10 . 15

Phe Pro Ser Trp Ser Cys Arg Arg Lys Phe Pro Val Val Lys Arg Tyr 20 25 30

Ser Tyr Arg Asn Ile Arg Phe Gly Leu Cys Ser Val Arg Ala Ser Gly 40 Gly Gly Ser Ser Gly Ser Glu Ser Cys Val Ala Val Arg Glu Asp Phe Ala Asp Glu Glu Asp Phe Val Lys Ala Gly Gly Ser Glu Ile Leu Phe Val Gln Met Gln Gln Asn Lys Asp Met Asp Glu Gln Ser Lys Leu Val Asp Lys Leu Pro Pro Ile Ser Ile Gly Asp Gly Ala Leu Asp His Val Val Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala Ala Glu Ser Ala 120 Lys Leu Gly Leu Lys Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Glu Asp Glu Phe Asn Asp Leu Gly Leu Gln Lys Cys Ile Glu His Val Trp Arg Glu Thr Ile Val Tyr Leu Asp Asp 170 Asp Lys Pro Ile Thr Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg Arg Leu Leu His Glu Glu Leu Leu Arg Arg Cys Val Glu Ser Gly Val Ser Tyr Leu Ser Ser Lys Val Asp Ser Ile Thr Glu Ala Ser Asp Gly Leu Arg Leu Val Ala Cys Asp Asp Asn Asn Val Ile Pro Cys Arg Leu Ala 235 Thr Val Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Gln Tyr Glu Val Gly Gly Pro Arg Val Cys Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Glu Asn Ser Pro Tyr Asp Pro Asp Gln Met Val Phe Met Asp Tyr Arg Asp Tyr Thr Asn Glu Lys Val Arg Ser Leu Glu Ala Glu Tyr Pro Thr Phe Leu Tyr Ala Met Pro Met Thr Lys Ser Arg Leu Phe Phe Glu 310 315 Glu Thr Cys Leu Ala Ser Lys Asp Val Met Pro Phe Asp Leu Leu Lys 330 Thr Lys Leu Met Leu Arg Leu Asp Thr Leu Gly Ile Arg Ile Leu Lys 345 Thr Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro

355 360 365

Asn Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val 370 375 380

His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro 385 390 395 400

Lys Tyr Ala Ser Val Ile Ala Glu Ile Leu Arg Glu Glu Thr Thr Lys 405 410 415

Gln Ile Asn Ser Asn Ile Ser Arg Gln Ala Trp Asp Thr Leu Trp Pro 420 425 430

Pro Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu Phe Gly Leu Ala Leu 435 440 445

Ile Val Gln Phe Asp Thr Glu Gly Ile Arg Ser Phe Phe Arg Thr Phe 450 455 460

Phe Arg Leu Pro Lys Trp Met Trp Gln Gly Phe Leu Gly Ser Thr Leu 465 470 475 480

Thr Ser Gly Asp Leu Val Leu Phe Ala Leu Tyr Met Phe Val Ile Ser 485 490 495

Pro Asn Asn Leu Arg Lys Gly Leu Ile Asn His Leu Ile Ser Asp Pro 500 505 510

Thr Gly Ala Thr Met Ile Lys Thr Tyr Leu Lys Val 515 520

<210> 22

<211> 1898

<212> DNA

<213> Adonis palaestina

<400> 22 AAAGGAGTGT TCTATTAATG TTACTGTCGC ATTCTTGCAA CACTTATATT CAAACTCCAT 60 TTTCTTCTTT TCTCTTCAAA ACAACAAACT AATGTGAGCA GAGTATCTGG CTATGGAACT 120 ACTTGGTGTT CGCAACCTCA TCTCTTCTTG CCCTGTGTGG ACTTTTGGAA CAAGAAACCT 180 TAGTAGTTCA AAACTAGCTT ATAACATACA TCGATATGGT TCTTCTTGTA GAGTAGATTT 240 TCAAGTGAGA GCTGATGGTG GAAGCGGGAG TAGAAGTTCT GTTGCTTATA AAGAGGGTTT 300 TGTGGATGAA GAGGATTTTA TCAAAGCTGG TGGTTCTGAG CTTTTGTTTG TCCAAATGCA 360 GCAAACAAG TCTATGGAGA AACAGGCCAA GCTCGCCGAT AAGTTGCCAC CAATACCTTT 420 TGGAGAATCC GTGATGGACT TGGTTGTAAT AGGTTGTGGA CCTGCTGGTC TTTCACTGGC 480 TGCAGAAGCT GCTAAGCTAG GGTTGAAAGT TGGCCTTATT GGTCCTGATC TTCCTTTTAC 540 AAATAATTAT GGTGTGGGG AAGACGAGTT CAAAGATCTT GGACTTGAAC GTTGTATCGA 600 GCATGCTTGG AAGGACACCA TCGTATATCT TGATAATGAT GCTCCTGTCC TTATTGGTCG 660 TGCATATGGA CGAGTTAGTC GACATTTGCT ACATGAGGAG TTGCTGAAAA GGTGTGTGGA 720

| | | | | | " >, | |
|------------|------------|------------|------------|------------|-------------|------|
| GTCAGGTGTA | TCATATCTTG | ATTCTAAAGT | GGAAAGGATC | ACTGAAGCTG | - | 780 |
| TAGCCTTGTA | GTTTGTGAAA | ATGAGATCTT | TATCCCTTGC | AGGCTTGCTA | CTGTTGCATC | 840 |
| TGGAGCAGCT | TCAGGGAAAC | TTTTGGAGTA | TGAAGTAGGT | GGCCCTCGTG | TTTGTGTCCA | 900 |
| AACCGCTTAT | GGGGTGGAGG | TTGAGGTGGA | GAACAATCCA | TACGATCCCA | ACTTAATGGT | 960 |
| ATTCATGGAC | TACAGAGACT | ATATGCAACA | GAAATTACAG | TGCTCGGAAG | AAGAATATCC | 1020 |
| AACATTTCTC | TATGTCATGC | CCATGTCGCC | AACAAGACTT | TTTTTTGAGG | AAACCTGTTT | 1080 |
| GGCCTCAAAA | GATGCCATGC | CATTCGATCT | ACTGAAGAGA | AAACTGATGT | CACGATTGAA | 1140 |
| GACTCTGGGT | ATCCAAGTTA | CAAAAGTTTA | TGAAGAGGAA | TGGTCATATA | TTCCTGTTGG | 1200 |
| TGGTTCTTTA | CCAAACACAG | AGCAAAAGAA | CCTAGCATTT | GGTGCTGCAG | CAAGCATGGT | 1260 |
| GCATCCAGCA | ACAGGCTATT | CGGTTGTACG | GTCACTGTCA | GAAGCTCCAA | AATATGCTTC | 1320 |
| TGTAATTGCA | AAGATTTTGA | AGCAAGATAA | CTCTGCGTAT | GTGGTTTCTG | GACAAAGTAG | 1380 |
| TGCAGTAAAC | ATTTCAATGC | AAGCATGGAG | CAGTCTTTGG | CCAAAGGAGC | GAAAACGTCA | 1440 |
| AAGAGCATTC | TTTCTTTTTG | GATTAGAGCT | TATTGTGCAG | CTAGATATTG | AAGCAACCAG | 1500 |
| AACATTCTTT | AGAACCTTCT | TCCGCTTGCC | AACTTGGATG | TGGTGGGGTT | TCCTTGGGTC | 1560 |
| TTCACTATCA | TCTTTCGATC | TCGTCTTGTT | TTCCATGTAC | ATGTTTGTTT | TGGCGCCAAA | 1620 |
| CAGCATGAGG | ATGTCACTTG | TGAGACATTT | GCTTTCAGAT | CCTTCTGGTG | CAGTTATGGT | 1680 |
| AAGAGCTTAC | CTCGAAAGGT | AGTCTCATCT | АТТАТТАААС | TCTAGTGTTT | CACCAAATAA | 1740 |
| ATGAGGATCC | TTCGAATGTG | TATATGATCA | TCTCTATGTA | TATCCTGTAC | TCTAATCTCA | 1800 |
| TAAAGTAAAT | GCCGGGTTTG | ATATTGTTGT | GTCAAACCGG | CCAATGATAT | AAAGTAAATT | 1860 |
| TATTGATACA | AAAGTAGTTT | TTTTCCTTAA | AAAAAAA | | | 1898 |

<210> 23

<400> 23

<211> 529

<212> PRT

<213> Adonis palaestina

Met Glu Leu Leu Gly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp 1 5 10 15

Thr Phe Gly Thr Arg Asn Leu Ser Ser Ser Lys Leu Ala Tyr Asn Ile 20 25 30

His Arg Tyr Gly Ser Ser Cys Arg Val Asp Phe Gln Val Arg Ala Asp 35 40 45

Gly Gly Ser Gly Ser Arg Ser Ser Val Ala Tyr Lys Glu Gly Phe Val 50 55 60

Asp Glu Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu Leu Leu Phe Val 65 70 75 80



| | | | | | | | | | ~ | | | | | | |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Gln | Met | Gln | Gln | Thr 85 | Lys | Ser | Met | Glu | Lys 90 | Gln | Ala | Lys | Leu | Ala 95 | Asp |
| Lys | Leu | Pro | Pro 100 | Ile | Pro | Phe | Gly | Glu 105 | Ser | .Val | Met | Asp | Leu 110 | Val | Val |
| Ile | Gly | Cys 115 | Gly | Pro | Ala | Gly | Leu 120 | Ser | Leu | Ala | Ala | Glu 125 | Ala | Ala | Lys |
| Leu | Gly 130 | Leu | Lys | Val | Gly | Leu 135 | Ile | Gly | Pro | Asp | Leu 140 | Pro | Phe | Thr | Asn |
| Asn 145 | Tyr | Gly | Val | Trp | Glu 150 | Asp | Glu | Phe | Lys | Asp 155 | Ļeu | Gly | Leu | Glu | Arg 160 |
| Cys | Ile | Glu | His | Ala 165 | Trp | Lys | Asp | Thr | Ile 170 | Val | Tyr | Leu | Asp | Asn 175 | Asp |
| Ala | Pro | Val | Leu 180 | Ile | Gly | Arg | Ala | Tyr 185 | Gly | Arg | Val | Ser | Arg 190 | His | Leu |
| Leu | His | Glu 195 | Glu | Leu | Leu | Lys | Arg 200 | Cys | Val, | Glu | Ser | Gly 205 | Val | Ser | Tyr |
| Leu | Asp 210 | Ser | Lys | Val | Glu | Arg 215 | Ile | Thr | Glu | Ala | Gly 220 | Asp | Gly | His | Ser |
| Leu 225 | Val | Val | Cys | Glu | Asn 230 | Glu | Ile | Phe | Ile | Pro 235 | Cys | Arg | Leu | Ala | Thr 240 |
| Val | Ala | Ser | Gly | Ala 245 | Ala | Ser | Gly | Lys | Leu 250 | Leu | Glu | Tyr | Glu | Val 255 | Gly |
| Gly | Pro | Arg | Val 260 | Cys | Val | Gln | Thr | Ala 265 | Tyr | Gly | Val | Glu | Val 270 | Glu | Val |
| Glu | Asn | Asn 275 | Pro | Tyr | Asp | Pro | Asn 280 | Leu | Met | Val | Phe | Met 285 | Asp | Tyr | Arg |
| Asp | Tyr 290 | Met | Gln | Gln | Lys | Leu 295 | Gln | Cys | Ser | Glu | Glu 300 | Glu | Tyr | Pro | Thr |
| Phe 305 | Leu | Tyr | Val | Met | Pro 310 | Met | Ser | Pro | Thr | Arg 315 | | Phe | Phe | Glu | Glu 320 |
| Thr | Cys | Leu | Ala | Ser 325 | Lys | Asp | Ala | Met | Pro 330 | Phe | Asp | Leu | Leu | Lys 335 | Arg |
| Lys | Leu | Met | Ser 340 | Arg | Leu | Lys | Thr | Leu 345 | Gly | Ile | Gln | Val | Thr 350 | Lys | Val |
| Tyr | Glu | Glu 355 | Glu | Trp | Ser | Tyr | Ile 360 | Pro | Val | Gly | Gly | Ser 365 | Leu | Pro | Asn |
| Thr | Glu 370 | Gln | Lys | Asn | Leu | Ala 375 | Phe | Gly | Ala | Ala | Ala 380 | Ser | Met | Val | His |
| Pro 385 | Ala | Thr | Gly | Tyr | Ser 390 | Val | Val | Arg | Ser | Leu 395 | Ser | Glu | Ala | Pro | Lys 400 |
| Tyr | Ala | Ser | Val | Ile | Ala | Lys | Ile | Leu | Lys | Gln | Asp | Asn | Ser | Ala | Tyr |

PCT/US99/12121

415

| wo 9 | 99/63055 | |
|------|----------|--|
|------|----------|--|

410

Val Val Ser Gly Gln Ser Ser Ala Val Asn Ile Ser Met Gln Ala Trp 420 425 430

Ser Ser Leu Trp Pro Lys Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu

Phe Gly Leu Glu Leu Ile Val Gln Leu Asp Ile Glu Ala Thr Arg Thr 450 455 460

Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp Trp Gly Phe 465 470 475 480

Leu Gly Ser Ser Leu Ser Ser Phe Asp Leu Val Leu Phe Ser Met Tyr 485 490 495

Met Phe Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His
500 505 510

Leu Leu Ser Asp Pro Ser Gly Ala Val Met Val Arg Ala Tyr Leu Glu 515 520 525

Arq

<210> 24

<211> 1370

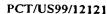
<212> DNA

<213> Potato

<400> 24

TAGCGGAGGA TGAGTTCAAA GATCTTGGTC TTCAAGCCTG CATTGAACAT GTTTGGCTGG 60 GATACCATTG TATATCTTGA TGATGATGAT CCTATTCTTA TTGGCCGTGC CTATGGAAGA 120 GTTAGTCGCC ATTTACTGCA CGAGGAGTTA CTCAAAAGGT GTGTGGAGGC AGGTGTTTTG 180 TATCTAAACT CGAAAGTGGA TAGGATTGTT GAGGCCACAA ATGGCCACAG TCTTGTAGAG 240 TGCGAGGGTG ATGTTGTGAT TCCCTGCAGG TTTGTGACTG TTGCATCGGG AGCAGCCTCG 300 GGGAAATTCT TGCAGTATGA GTTGGGAGGT CCTAGAGTTT CTGTTCAAAC AGCTTATGGA 360 GTGGAAGTTG AGGTCGATAA CAATCCATTT GACCCGAGCC TGATGGTTTT CATGGATTAT 420 AGAGACTATG TCAGACACGA CGCTCAATCT TTAGAAGCTA AATATCCAAC ATTTCTCTAT 480 GCCATGCCCA TGTCTCCAAC ACGAGTCTTT TTCGAGGAAA CTTGTTTGGC TTCAAAAGAT 540 GCAATGCCAT TCGATCTGTT AAAGAAAAAA TTGATGTTAC GATTGAACAC CCTCGGTGTA 600 AGAATTAAAG AAATTTATGA GGAGGAATGG TCTTACATAC CAGTTGGAGG ATCTTTGCCA 660 AATACAGAAC AAAAAACACT TGCATTTGGT GCTGCTGCTA GCATGGTTCA TCCAGCCACA 720 GGTTATTCAG TCGTCAGATC ACTGTCTGAA GCTCCAAAAT GCGCCTTCGT GCTTGCAAAT 780 ATATTACGAC AAAATCATAG CAAGAATATG CTTACTAGTT CAAGTACCCC GAGTATTTCA 840 ACTCAAGCTT GGAACACTCT TTGGCCACAA GAACGAAAAC GACAAAGATC GTTTTTCCTA 900

WO 99/63055



| | | | | | • | |
|---|------------|------------|------------|------------|------------|------|
| TTTGGACTGG | CTCTGATATT | GCAGCTGGAT | ATTGAGGGGA | TAAGGTCATT | TTTCCGCGCG | 960 |
| TTCTTCCGTG | TGCCAAAATG | GATGTGGCAG | GGATTTCTTG | GTTCAAGTCT | TTCTTAGCAG | 1020 |
| ACCTCATGTT | ATTTGCCTTC | TACATGTTTA | TTATTGCACC | AAATGACATG | AGAAGAGGCT | 1080 |
| TAATCAGACA | TCTTTTATCT | GATCCTACTG | GTGCAACATT | GATAAGAACT | TATCTTACAT | 1140 |
| TTTAGAGTAA | ATTCCTCCTA | CAATAGTTGT | TGAAAGAGGC | CTCATTACTT | CAGATTCATA | 1200 |
| ACAGAAATCG | CGGTCTCTCG | AGGCCTTGTA | TATAACATTT | TCACTAGGTT | AATATTGCTT | 1260 |
| GAATAAGTTG | CACAGTTTCA | GTTTTTGTAT | CTGCTTCTTT | TTTGTCCAAG | ATCATGTATT | 1320 |
| GACCAATTTA | TATACATTGC | CAGTATATAT | AAATTTTATA | AAAAAAAA | | 1370 |
| <210> 25 <211> 377 <212> PRT <213> Potat | :0 | | | | · | |

<400> 25

Asp Glu Phe Lys Asp Leu Gly Leu Gln Ala Cys Ile Glu His Val Trp

Arg Asp Thr Ile Val. Tyr Leu Asp Asp Asp Pro Ile Leu Ile Gly

Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu His Glu Glu Leu Leu

Lys Arg Cys Val Glu Ala Gly Val Leu Tyr Leu Asn Ser Lys Val Asp

Arg Ile Val Glu Ala Thr Asn Gly His Ser Leu Val Glu Cys Glu Gly

Asp Val Val Ile Pro Cys Arg Phe Val Thr Val Ala Ser Gly Ala Ala

Ser Gly Lys Phe Leu Gln Tyr Glu Leu Gly Gly Pro Arg Val Ser Val

Gln Thr Ala Tyr Gly Val Glu Val Glu Val Asp Asn Asn Pro Phe Asp 120

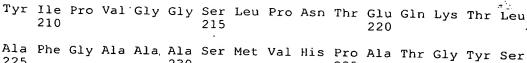
Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp Tyr Val Arg His Asp 135

Ala Gln Ser Leu Glu Ala Lys Tyr Pro Thr Phe Leu Tyr Ala Met Pro 150 155 -

Met Ser Pro Thr Arg Val Phe Phe Glu Glu Thr Cys Leu Ala Ser Lys 170

Asp Ala Met Pro Phe Asp Leu Leu Lys Lys Leu Met Leu Arg Leu 180

Asn Thr Leu Gly Val Arg Ile Lys Glu Ile Tyr Glu Glu Glu Trp Ser 200



Val Val Arg Ser Leu Ser Glu Ala Pro Lys Cys Ala Phe Val Leu Ala

Asn Ile Leu Arg Gln Asn His Ser Lys Asn Met Leu Thr Ser Ser Ser

Thr Pro Ser Ile Ser Thr Gln Ala Trp Asn Thr Leu Trp Pro Gln Glu

Arg Lys Arg Gln Arg Ser Phe Phe Leu Phe Gly Leu Ala Leu Ile Leu 295

Gln Leu Asp Ile Glu Gly Ile Arg Ser Phe Phe Arg Ala Phe Phe Arg 315

Val Pro Lys Met Met Trp Gly Phe Leu Gly Ser Ser Leu Ser Xaa Ala

Asp Leu Met Leu Phe Ala Phe Tyr Met Phe Ile Ile Ala Pro Asn Asp 345

Met Arg Arg Gly Leu Ile Arg His Leu Leu Ser Asp Pro Thr Gly Ala 360

Thr Leu Ile Arg Thr Tyr Leu Thr Phe

<210> 26

<211> 533

<212> PRT

<213> Chimeric lettuce/potato

Met Glu Cys Phe Gly Ala Arg Asn Met Thr Ala Thr Met Ala Val Phe 10

Thr Cys Pro Arg Phe Thr Asp Cys Asn Ile Arg His Lys Phe Ser Leu

Leu Lys Gly Arg Arg Phe Thr Asn Leu Ser Ala Ser Ser Ser Leu Arg

Gln Ile Lys Cys Ser Ala Lys Ser Asp Arg Cys Val Val Asp Lys Gln

Gly Ile Ser Val Ala Asp Glu Glu Asp Tyr Val Lys Ala Gly Gly Ser

Glu Leu Phe Phe Val Gln Met Gln Arg Thr Lys Ser Met Glu Ser Gln

Ser Lys Leu Ser Glu Lys Leu Ala Gln Ile Pro Ile Gly Asn Cys Ile

Leu Asp Leu Val Val Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala



بن جر 115 120 125 Ala Glu Ser Ala Lys Leu Gly Leu Asn Val Gly Leu Ile Gly Pro Asp 135 Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Gln Asp Glu Phe Ile Gly 150 Leu Gly Leu Glu Gly Cys Ile Glu His Ser Trp Lys Asp Thr Leu Val Tyr Leu Asp Asp Ala Asp Pro Ile Arg Ile Gly Arg Ala Tyr Gly Arg Val His Arg Asp Leu Leu His Glu Glu Leu Leu Arg Arg Cys Val Glu 200 Ser Gly Val Ser Tyr Leu Ser Ser Lys Val Glu Arg Ile Thr Glu Ala Pro Asn Gly Tyr Ser Leu Ile Glu Cys Glu Gly Asn Ile Thr Ile Pro Cys Arg Leu Ala Thr Val Ala Ser Gly Ala Ala Ser Gly Lys Phe Leu Glu Tyr Glu Leu Gly Gly Pro Arg Val Ser Val Gln Thr Ala Tyr Gly 265 Val Glu Val Glu Val Asp Asn Asn Pro Phe Asp Pro Ser Leu Met Val 280 Phe Met Asp Tyr Arg Asp Tyr Val Arg His Asp Ala Gln Ser Leu Glu Ala Lys Tyr Pro Thr Phe Leu Tyr Ala Met Pro Met Ser Pro Thr Arg 310 Val Phe Phe Glu Glu Thr Cys Leu Ala Ser Lys Asp Ala Met Pro Phe Asp Leu Leu Lys Lys Leu Met Leu Arg Leu Asn Thr Leu Gly Val 345 Arg Ile Lys Glu Ile Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly 360 Gly Ser Leu Pro Asn Thr Glu Gln Lys Thr Leu Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu 390 395 Ser Glu Ala Pro Lys Cys Ala Phe Val Leu Ala Asn Ile Leu Arg Gln Asn His Ser Lys Asn Met Leu Thr Ser Ser Ser Thr Pro Ser Ile Ser 425 Thr Gln Ala Trp Asn Thr Leu Trp Pro Gln Glu Arg Lys Arg Gln Arg



- Ser Phe Phe Leu Phe Gly Leu Ala Leu Ile Leu Gln Leu Asp Ile Glu
 450 455 460
- Gly Ile Arg Ser Phe Phe Arg Ala Phe Phe Arg Val Pro Lys Trp Met 465 470 475 480
- Trp Gln Gly Phe Leu Gly Ser Ser Leu Ser Xaa Ala Asp Leu Met Leu 485 490 495
- Phe Ala Phe Tyr Met Phe Ile Ile Ala Pro Asn Asp Met Arg Arg Gly 500 505 510
- Leu Ile Arg His Leu Leu Ser Asp Pro Thr Gly Ala Thr Leu Ile Arg 515 520 525
- Thr Tyr Leu Thr Phe 530
- <210> 27
- <211> 374
- <212> PRT
- <213> Arabidopsis thaliana
- <400> 27
- Glu Asp Glu Phe Asn Asp Leu Gly Leu Gln Lys Cys Ile Glu His Val 1 5 10 15
- Trp Arg Glu Thr Ile Val Tyr Leu Asp Asp Asp Lys Pro Ile Thr Ile 20 25 30
- Gly Arg Ala Tyr Gly Arg Val Ser Arg Arg Leu Leu His Glu Glu Leu 35 40 45
- Leu Arg Arg Cys Val Glu Ser Gly Val Ser Tyr Leu Ser Ser Lys Val
- Asp Ser Ile Thr Glu Ala Ser Asp Gly Leu Arg Leu Val Ala Cys Asp 65 70 75 80
- Asp Asn Asn Val Ile Pro Cys Arg Leu Ala Thr Val Ala Ser Gly Ala 85 90 95
- Ala Ser Gly Lys Leu Leu Gln Tyr Glu Val Gly Gly Pro Arg Val Cys 100 105 110
- Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Glu Asn Ser Pro Tyr 115 120 125
- Asp Pro Asp Gln Met Val Phe Met Asp Tyr Arg Asp Tyr Thr Asn Glu 130 135 140
- Lys Val Arg Ser Leu Glu Ala Glu Tyr Pro Thr Phe Leu Tyr Ala Met 145 150 155 160
- Pro Met Thr Lys Ser Arg Leu Phe Phe Glu Glu Thr Cys Leu Ala Ser 165 170 175
- Lys Asp Val Met Pro Phe Asp Leu Leu Lys Thr Lys Leu Met Leu Arg
- Leu Asp Thr Leu Gly Ile Arg Ile Leu Lys Thr Tyr Glu Glu Glu Trp



| 200 | 205 | े ज हैं |
|-----|-----|---------|

Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Asn 215

Leu Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr 235

Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys Tyr Ala Ser Val Ile 250

Ala Glu Ile Leu Arg Glu Glu Thr Thr Lys Gln Ile Asn Ser Asn Ile 260 265

Ser Arg Gln Ala Trp Asp Thr Leu Trp Pro Pro Glu Arg Lys Arg Gln

Arg Ala Phe Phe Leu Phe Gly Leu Ala Leu Ile Val Gln Phe Asp Thr 295

Glu Gly Ile Arg Ser Phe Phe Arg Thr Phe Phe Arg Leu Pro Lys Trp 310 315

Met Trp Gln Gly Phe Leu Gly Ser Thr Leu Thr Ser Gly Asp Leu Val 330

Leu Phe Ala Leu Tyr Met Phe Val Ile Ser Pro Asn Asn Leu Arg Lys

Gly Leu Ile Asn His Leu Ile Ser Asp Pro Thr Gly Ala Thr Met Ile 360

Lys Thr Tyr Leu Lys Val 370

<210> 28

<211> 1002

<212> DNA

<213> Adonis palaestina

<400> 28

ATTCATCTTC AGCAGCGCTG TCGTACTCTT TCTATATCTT CTTCCATCAC TAACAGTAGT 60 CGCCGACGGT TGAATCGGCT ATTCGCCTCA ACGTCAACTA TGGGTGAAGT CACTGATGCT 120 GGAATGGATG CTGTTCAGAA GCGGCTCATG TTCGACGACG AATGTATTTT GGTGGATGAG 180 AATGACAAGG TCGTCGGGCA TGATTCCAAA TACAACTGTC ATTTGATGGA AAAGATAGAG 240 GCAGAAAATT TGCTTCACAG AGCCTTCAGT GTTTTCTTGT TCAACTCAAA ATATGAATTG 300 CTTCTTCAGC AACGATCCGC CACAAAGGTA ACATTCCCGC TCGTATGGAC AAACACATGT 360 TGCAGTCATC CTCTCTTCG TGATTCCGAG CTCATAGAAG AAAATTATCT CGGTGTACGA 420 AACGCTGCAC AAAGAAAGCT TTTAGACGAG CTAGGCATTC CAGCTGAAGA TGTCCCAGTT 480 GATGAATTTA CTCCTCTTGG TCGCATTCTT TACAAAGCTC CATCTGACGG CAAATGGGGA 540 GAGCACGAAT TGGACTATCT CCTATTTATT GTCCGAGATG TGAAATACGA TCCAAACCCA 600

WO 99/63055 PCT/US99/12121

| GATGAAGTTG | CTGATGCTAA | GTATGTTAAT | CGCGAGGAGT | TGAGAGAGAT | ACTGAGAAAA | 660 |
|--|--------------|-------------|------------|------------|------------|------|
| GCTGATGCTG | GTGAAGAGGG | ACTCAAGTTG | TCTCCTTGGT | TTAGATTGGT | TGTTGATAAC | 720 |
| TTTTTGTTCA | AGTGGTGGGA | TCATGTAGAG | CAGGGTACGA | TTAAGGAAGT | TGCTGACATG | 780 |
| AAAACTATCC | ACAAGTTGAC | TTAAGAGGAC | TTCTCTCCTC | TGTTCTACTA | TTTGTTTTTT | 840 |
| GCTACAATAA | GTGGGTGGTG | ATAAGCAGTT | TTTCTGTTTT | СТТТААТТТА | TGGCTTTTGA | 900 |
| ATTTGCCTCG | ATGTTGAACT | TGTAACATAT | TTAGACAAAT | ATGAGACCTT | GTAAGTTGAA | 960 |
| TTTGAGGCTG | AATTTATATT | TTTGGGAACA | TAATAATGTT | AA | | 1002 |
| <210> 29 <211> 1271 <212> DNA <213> Adon: | is palaestir | na | | | | |
| <400> 29 | ርፕሞጥርርርጥርር | מררמררמידרמ | AACCCACCCA | AATTTCTCTG | ma | 60 |
| | | | | CCTTGTTTAC | | 60 |
| | | | | GGCCCCTTTC | | 120 |
| | | | | TCTTCTCCAC | | 180 |
| | | | | | | 240 |
| | | | | AGCAGCGTTG | | 300 |
| | | | | TTGAATCGGC | | 360 |
| | | | | GCCGTCCAGA | | 420 |
| | | | | GTCGTCGGAC | | 480 |
| | | | | TTGCTTCACA | , | 540 |
| | | | | CAACGATCTG | | 600 |
| | | | | CCCCTCTTCC | | 660 |
| | | | | CAAAGGAAGC | | 720 |
| | | | • | ACTCCTCTTG | | 780 |
| | | | | CTGGACTATC | | 840 |
| | | | | GCTGACGCTA | | 900 |
| | | | | GGTGAAGAGG | | 960 |
| | | | | AAGTGGTGGG | | 1020 |
| | | | | CACAAGTTGA | | 1080 |
| | | • | | TGGATGGTGA | | 1140 |
| | | | | TGTTGAACTT | | 1200 |
| TAGTCAAATA | TGAGACCTTG | TGAGTTGAAT | TTGAGGTTAT | ATTTATAGTT | TTGGGAACAT | 1260 |
| | | | | | | |

PCT/US99/12121

| WO 99/63055 | |
|-------------|--|
| | |

| ААААААААА | in the factor of | 1271 |
|--|--|------|
| <210> 30 <211> 1109 <212> DNA <213> Haematococcus pluvialis | | |
| <400> 30 TGGAACCTGG CCCGGCGCA GTCCGATGCC GCGATGCTTC GTTCGTTGCT | C. C. C. C. C. C. C. C. C. C. C. C. C. C | |
| ACGCATATCC CGCGCGTGAA CTCCGCCCAG CAGCCCAGCT GTGCACACGC | | 60 |
| | | 120 |
| TTTAAGCTCA GGAGCATGCA GCTGCTTGCC GAGGACCGCA CAGACCACAT | | 180 |
| AGCACCTGGG CAGGCGGCA GTCGCAGGAT GAGCTGATGC TGAAGGACGA | | 240 |
| GTGGATGCTG ACGACAACAT CACAGGCCAT GCCAGCAAGC TGGAGTGCCA | CAAATTCCTA | 300 |
| CCACATCAGC CTGCAGGCCT GCTGCACCGG GCCTTCTCTG TGTTCCTGTT | TGACGACCAG | 360 |
| GGGCGACTGC TGCTGCAACA GCGTGCACGC TCAAAAATCA CCTTCCCAAG | TGTGTGGACG | 420 |
| AACACCTGCT GCAGCCACCC TCTACATGGG CAGACCCCAG ATGAGGTGGA | CCAACTAAGC | 480 |
| CAGGTGGCCG ACGGCACAGT ACCTGGCGCA AAAGCTGCTG CCATCCGCAA | GTTGGAGCAC | 540 |
| GAGCTGGGGA TACCAGCGCA CCAGCTGCCG GCAAGCGCGT TTCGCTTCCT | CACGCGTTTG | 600 |
| CACTACTGTG CCGCGGACGT GCAGCCGGCT GCGACACAAT CAGCGCTCTG | GGGCGAGCAC | 660. |
| GAGATGGACT ACATCTTATT CATCCGGGCC AACGTCACCT TGGCGCCCAA | CCCTGACGAG | 720 |
| GTGGACGAAG TCAGGTACGT GACGCAAGAG GAGCTGCGGC AGATGATGCA | GCCGGACAAC | 780 |
| GGGTTGCAAT GGTCGCCGTG GTTTCGCATC ATCGCCGCGC GCTTCCTTGA | GCGTTGGTGG | 840 |
| GCTGACCTGG ACGCGGCCCT AAACACTGAC AAACACGAGG ATTGGGGAAC | GGTGCATCAC | 900 |
| ATCAACGAAG CGTGAAGGCA GAAGCTGCAG GATGTGAAGA CACGTCATGG | GGTGGAATTG | 960 |
| CGTACTTGGC AGCTTCGTAT CTCCTTTTTC TGAGACTGAA CCTGCAGAGC | TAGAGTCAAT | 1020 |
| GGTGCATCAT ATTCATCGTC TCTCTTTTGT TTTAGACTAA TCTGTAGCTA | GAGTCACTGA | 1080 |
| TGAATCCTTT ACAACTTTCA AAAAAAAA | | 1109 |
| <210> 31 <211> 985 <212> DNA <213> Lactuca sativa | | · |
| <400> 31 TGCCAAAATG TTGAAATTTC CCCCTTTTAA AACCATTGCT ACCATGATCT | CTTCTCCATA | 60 |
| TTCTTCCTTC TTGCTGCCTC GGAAATCTTC TTTCCCTCCA ATGCCGTCTC | TCGCAGCCGC | 120 |
| TAGTGTTTTC CTCCACCCTC TTTCGTCTGC CGCTATGGGC GATTCCAGCA | TGGATGCTGT | 180 |
| CCAGCGACGT CTCATGTTCG ATGACGAATG CATTTTGGTG GATGAGAATG | ACAAAGTGGT | 240 |
| TGGCCATGAT ACTAAATACA ATTGTCATTT GATGGAGAAG ATTGAAAAGG | GAAATATGCT | 300 |

| WO 99/63055 | |
|-------------|--|



| ACACAGAGGA TTCAGTGTGT TCTTGTTCAA CTCGAAATAT GAATTACTCC TTCAGCAACG TTCTGCAACC AAGGTGACTT TCCCTTTGGT ATGACAACA ACGTGTTGCA GCCATCCACT 420 ATACAGGGAG AGGAGCTTA TTGACGAAAA CGCCCTTGGG GTGAGGAATG CTGCACAGAG 480 GAAGCTCCTG GATGAACTCG GCATCCCTG AGCAGTGTT CCGGTGATG AGTTCACTCC 540 ATTGGGTGCC ATTCTATACA AGGCCGCATC GGATGGAAGA TGGGGGAGAC ATGAACTTGA 600 TTACCTGCTG TTTATGGTAC GTGATGTGG TTTGGATCG AACCCAGATG AAGTGAAAGA 660 TGTAAAATAT GTGAACCGG AAGACCTAA GGACTGTAA AGGAAGGGG ATGCTGGTGA 720 AGAGGGTGTG AAGCCGGC AAGACCCTAAC CGAACTTT GATATCAAAA CAATCCACAA 840 ACTCACATAA AAACACTAC CTAGTAGGAG AGAGCATTT GATATCAAAA CAATCCACAA 840 ACTCACATAA AAACACTAC CTAGTAGGA AGAGCATTT TTGGATACT TGTTTATGTT 900 GAAATTGAAA TTCAGATGAA TGCTTGTATT TATTTCTATT TGGACAAACT TCAACTTCTT 966 ATTTGCTACCT TATCAGAAAA AAAAA 988 <2210 | | | | | | | |
|--|------------------------|----------------|------------|---------------|------------|------------|-----|
| ATACAGGGAG AGTGAGCTTA TTGACGAAAA CGCCCTTGGG GTGAGGAATG CTGCACAGAG (AGAGCTCCTG GATGAACTCG GATGAACTCG GATGAACTCG AGCAGATGTT CGGGTTGATG AGTTCACTCC (ATTGAGCTCG ATCTATACA AGGCCGCATC GGATGGAAGG TGGGGAGAAC ATGAACTTGA (AGAGCTGGA ATTGACTTGA CAGAGCTGGA AGGCCGATG TTTATGGTAC GTGAGTGTGG TTGACACTGG AAGAGCTGGA AAGAGCTGGA AAGAGCTGGA AAGAGCGGA ATGAACTGA (AGAGCGGA AGAGCGGA AGAGCGGA AGAGCGGA AGAGCGGAAC ATGAACTGAC (AGAGCGGA AGAGCGGA AC CGAGAGTTAT ATGATATACA CAATCCACAAA AACACTACA CTAGTAGGAA AGAGCATAT ATGAGAAAA CAATCCACAAA ACACTACA CTAGTAGGAA AGAGCGATTAT ATGAGAAAAT TGATATATGT (AGAATTGAAAA TTCAGATGAAA AAAAA (AGATCACAA AAAAA (AGATCACAA AAAAA (AGATCACAA AAAAA (AGATCACAA AAAAA (AGATCACAA AAAAA (AGATCACAA AAAAA (AGATCACA AAAAA (AGATCACA AAAAA (AGATCACA AAAAA (AGATCACA AAAAA (AGATCACA AAAAA (AGATCACA AAAAACCTTCT CACTTAACCG CGCCATGGG TGACGATTCC GGCATGAGCG (ACTTCACC ACCATAACCG CCGCCATGGG TGACGATTCC GGCATGACCG (ACTTCACC ACCATAACCG CCGCCATGGG TGACGATTCC GGCATGACCA (ACCATAACCA ACCATAACCA CCGCCATGGG TGACGATTCC GGCATGACAA AATGACAATG (ACCATCACAA ACCATATGTC ACTTGATGAA AATGCAATTT CTCTTCAGC (ACCTTCACC ACCATAACCAA ACCATATGTC ACTTGATGAA AAACACTATA CTCCTTCAGC (ACCTTCACC ACCATAACCAA ACCATATGTC ACTTGATGAA AAACACTATA CTCCTTCAGC (ACCTTCATGA AACCATGAA AACCAATGAA AACCAATGTC TTGATGAAA AACCAATGAA AACCAATGAA AACCAATGAA AACCAATGAA AACCAATGAA AACCAATGAA AACCAATGAA AACCAATGAA AACCAATGAA AACCAATGAA AACCAATGAA AACCAATGAA AACCAATGAA AACCAATGAA AACCAATGAA CCTGCGAAA AACCAAGGAG CTAAATTCCC CTCCTGAAAA AAACACCTT TGCAGCCAA AACCACCT GAGAGAACC TACAAGAGAA CCTGCAACAA AACCAATGAA CCAAACCCT GAGAGAACCA CTCTGAGAAA AACCAATGAA AACCAATGAA CACCTGAAGAGAC TTCAAAAGAACA TAAATTGCT CAACAGGAC CTCTGAAGAACC TTCTGAGAAAAAACCATTC AAAAGGAGA TTAAAAAGAATT ACAAAGAACA TAAATTGCTG AAAACCATTC AAAAGGAGAC TTAAAAAAAACAATT ACAAAGAACAA TAAATTGTG AAAACCATTC AAAAGGAGAC TTAAAAAAAACCATTC TAAAAAAAACAATTA ATGAAAAAACCATTC AAAAGGAGAC TTAAAAAAAACAATTA ATTAAAAAAAACAATTA ATTAAAAAAAA | ACACAGAGCA | TTCAGTGTGT | TCTTGTTCAA | CTCGAAATAT | GAATTACTCC | TTCAGCAACG | 360 |
| ATTGGGTCCC GATGAACTCG GCATCCCTGG AGCAGATGTT CCGGTTGATG AGTTCACTCC ATTGGGTCCC ATTCTATACA AGGCCCCATC GGATGGAAAG TGGGGAGAAC ATGAACTTGA ATTGGGTCCC ATTCTATACA AGGCCCGATC GGATGGAAAG TGGGGAGAAC ATGAACTTGA TTACCTGCTG TTTATGGTAC GTGATGTTGG TTTGGATCCG AACCCAGATG AAGTGAAAGAA AGAGGTGAAA AGAGCTGCAC GGAATTGGTA AGGAAGAGCGAA AGAGCTGAA AGAGCTGAAA AGAGGTGGTG AAGCCTGCC CGTGGTTCAA ATTGATGTC GATAATTCT TGTTTCAGTG 780 GTGGGATCGA CTCCATAAGG GAACCCTAAC CGAAGCTATT GATATGAAAA CAATCCACAA 840 ACTCACATAA AAACACTACA CTAGTAGGAA AGAGGATTAT ATGAGATATT TGTTATATGT 900 GAAATTGAAA TTCAGAAAAA AAAAA CCIAC TATCAGAAAAA AAAAA C210> 32 C11> 988 C212> DNA C211> 988 C212> DNA C213> Lactuca sativa C400> 32 TATTCGCTTC AAAATCCTCT CCATTAACTG CTCAAATCTC CACCTTCGCC GGTCTTAATC TCCGCCCGGCG CACTTTCACC ACCATAACCG CCGCCATGGG TGACGATTC GGCATGGACG 120 CTGTCCAGAG ACGTCTCATG TTTGATGATG AATGCAATTT GGTGATGAA AATGACAATG 180 TTCTTGGGCA TGATACCAAA TACAATTGTC ACTTGATTC GGTGATGAA AATGACAATG 180 TCCTTCATGA AGCATTCAGT GTATTTTAT TCAATTCAAA ATACGAATTA CTCCTTCAGC 300 AAAAGGTCAGA AACCAAGGTG ACATTTCTT TGGTATGGA GAAGATTGAA AATGACAATG 180 TCCTTCATAG AGCATTCAGT GTATTTTAT TCAATTCAAA ATACGAATTA CTCCTTCAGC 300 AAAAGGTCAGA AACCAAGGTG ACATTTCCTT TGGTATGGA GAAGATTGAA AATGCACATC 360 CACTATACAG AGAATCGGAG TTAATTCCC ACATTGAGA AAACACCTGT TGCAGCCATC 360 CACATTAACAG AGAATCGGAG TTAATTCCC ACATTGAGGA TGTTCCAGTT GATGAGTTCA 480 CACATTAACAG AGAATCGGAG TTAATTCCC ACATTGAGGA TGTTCCAGTT GATGAGTTCA 480 CACATTAACAG AGAATCGGAG TTAATTCCC ACTTGATGGA CAAACACCTT TGCAGCCATC 360 CACATTAACAG AGAATCGGAG TTAATTCCC ACTTGATGGA CAACACCCT GATGAGTTCA 480 CACATTAACAG AGAATCGGAG TTAATTCCC TGCGTGAAGA TGTTCCAGTT GATGAGTTCA 480 CACATTAACAG AGAATCGGAG TTAATTCCC TGCGTGAGA CCCTAAACCCT GATGAGTTCA 480 CACATTAACA ATACGTGAAC CAAGAAGCT TAAAAGAGCT ACTAAAGGAG GCGATGCGG 660 CACATTAACAA ATACCTACA AAGAGGCT TAAAAGAGCT ACTAAAGAGC TATTGTTCA 720 AATGGTGGAG TTAAATTC TCCCATGGT TTAAAGAGCT ACTAAAGAGC AATACCCT GATGAGGTGT 780 AATGGTGGAG TCTGAAAATG CAAGAGAGT TAAAAGAGCT ATAATTTTT TTTCTTTTT TTCTTTTTT TGGGAGGTGT 780 AATGGTGGGA TCAAGAAAAAG GTTAAATTT ATGGGGA | TTCTGCAACC | AAGGTGACTT | TCCCTTTGGT | ATGGACAAAC | ACGTGTTGCA | GCCATCCACT | 420 |
| ATTGGGTCGC ATTCTATACA AGGCCGCATC GGATGGAAAG TGGGGAGAAC ATGAACTTGA 600 TTACCTGCTG TTTATGGTAC GTGATGTTGG TTTGGATCCG AACCCAGATG AAGTGAAAGA TGTAAAAATAT GTGAACCGGG AAGAGCTGAA GGAATTGGTA AGGAAGGCGG ATGCTGGTGA 720 AGAGGGTGTG AAGCTGTCCC CGTGGTTCAA ATTGATTGTC GATAATTCT TGTTTCAGTG 780 GTGGGATCGA CTCCATAAGG GAACCCTAAC CGAAGCTATT GATATGAAAA CAATCCACAA 840 ACTCACATAA AAACACTACA CTAGTAGGAG AGAGGATTAT ATGAGATATT TGTTATATGT 900 GAAAATTGAAA TCAGAAAAA AAAAA **C210> 32 C211> 988 <*212> DNA <*213> Lactuca sativa **C400> 32 TATTCGCTTC AAAATCCTT CCATTAACTG CTCAAATCTC CACCTTCGCC GGTCTTAATC CCGCCGGGG CACTTTCACC ACCATAACCG CCGCCATGGG TGACGATTCA GGCATGGACG 120 CTGTCCAGAG ACGTCTCATG TTTGATGATG AATGCATTT GGTTGATGAA AAACACTACA TCAGATGAT TCAGATGAA 180 TCCTTCGGCA TGATACCAAA TACAATTGTC ACTGATGGA GAAGATTGAG AAAGATAATT 240 TGCTTCATAG AGCATCAGT GTATTTTATT TCAATTCAAA ATACGAATTA CTCCTTCAGC 300 AAAGGTCAGA AACCAAGGTG ACATTTCCTT TGGTATGGA GAAGATTGAG AAAGATAATT 240 TGCTTCATAGA AACCAAGGTG ACATTTCCTT TGGTATGGA GAAGATTGAG AAAGATAATT 240 AAAGGTCAGA AACCAAGGTG ACATTTCCTT TGGTATGGA AAACACCTGT TGCAGCCATC 360 CACTATACAG AGAATCGGA TTAATTCCC AAAATCCCC TGGGGTCAGA AATGCTGCAC 420 AGAGGAAGCT TCTAGATGAA CCGGTATCC CTGCTGAAGA TGTTCCAGTT GATGAGTTCA 480 CACATTTAGG TCGCATGTTG TACAAGGCTC CATCTGATGG AAAAACCCTGT TGCAGCCATC 360 CACACTTTAGG TCGCATGTTG TACAAGGCTC CATCTGATGG AAAATGCCCT TGGGGTCAGA AATGCTGCAC 420 AGAGGAAGCT TCTAGATGAA CTCGGTATCC CTGCTGAAGA TGTTCCAGTT GATGAGTTCA 480 CAACTTTAGG TCGCATGTTC TACAAGGCTC CATCTGATGG AAAATGCCCT TGAGAGCTGG GACAATGAAG 540 CTGGACATTAG ATACCTACC CAAGAGAGT TAAAAGCCTT TGGGGTCAGA CTCTTGTTCA 720 AATGGTGGG TTTGAAATTG TCCCCATGGT TAAAAGAGTT ACTAAGGAAC GCGAATCCGG 660 CGGACATTAG ATACGTGAAC CAAGAAGAGT TAAAAGAGTT ACTAAGGAAC TTCTTGTTCA 720 AATGGTGGGA TCATGCCAA AAGGGGACC TCAATGAAGC AATTGACATG AAAACCATTC 780 AATGGTGGGA TCATGTCCAA AAGGGGACC TCAATGAAGC AATTGACATG AAAACCATTC 780 AATGGTGGGA TCATGTCCAA AAGGGGACC TCAATGAAGC AATTGACATG AAAACCATTC 780 ATAGGTGGA TCATGTCCAA AAGGGGACC TCAATGAAGC AATTGAACCA TTCTTGTTCA 720 AATGGTGGGA TCATGTCCAA AAGGGGACC TCAATGAAGC A | ATACAGGGAG | AGTGAGCTTA | TTGACGAAAA | CGCCCTTGGG | GTGAGGAATG | CTGCACAGAG | 480 |
| TTACCTECTE TITATEGTAC GTGATGTTG TTTGGATCCG AACCCAGATG AAGTGAAAGA 660 TGTAAAAATAT GTGAACCGGG AAGAGCTGAA GGAATTGGTA AGGAAGGCGG ATGCTGGTGA 720 AGAGGGTGTG AAGCTGTCCC CGTGGTTCAA ATTGATTGTC GATAATTTCT TGTTTCAGTG 780 GTGGGATCGA CTCCATAAGG GAACCCTAAC CGAAGCTATT GATATGAAAA CAATCCACAA 840 ACTCACATAA AAACACTACA CTAGTAGGAG AGAGGATTAT ATGAGATATT TGTTATATGT 900 GAAATTGAAA TTCAGATGAA TGCTTGTATT TATTTCTATT TGGACAAACT TCAACTTCTT 960 TTTGCTACCT TATCAGAAAA AAAAA 9885 <211> 988 <2121> DNA <211> 988 <212> DNA <213> Lactuca sativa <400> 32 TATTCGCTTC AAAATCTCTT CCATTAACTG CTCAAATCTC CACCTTCGCC GGTCTTAATC 60 CCGCCGGGG CACTTTCACC ACCATAACCG CCGCCATGGG TGACGATTC GGCATGGACG 120 CTGTCCAGAG ACGTCTCATG TTTGATGATG AATGCATTTT GGTTAGATAA AAAGATCAATG 180 TTCTTGGGCA TGATACCAAA TACAATTGTC ACTGATGGA GAAGATTGAG AAAGATAATT 240 TGCTTCATAG AGCATTCAGT GTATTTTAT TCAATTCAAA ATACGAATTA CTCCTTCAGC 300 AAAGGTCAGA AACCAAGGTG ACATTTCCTT TGGTATGGA AAACACCTGT TGCAGCCATC 360 CACTATACAG AGAATCGGA TTAATCCCG AAAATCCCT TGGGGTCAGA AATGCACATC 360 CACACTATACAG AGAATCGGA TTAATCCCG AAAATCCCT TGGGGTCAGA AATGCACATC 360 CACACTATACAG AGAATCGGA TTAATCCCG AAAATCCCT TGGGGTCAGA AATGCTGCAC 420 AGAGGAAGCT TCTAGATGAA CTCGGTATCC CACCTGAAGA TGTTCCAGGTT GATAGCTAC 480 CACACTATACAG AGAATCGGA TTAATCCCG AAAATGCCCT TGGGGTCAGA AATGCTGCAC 420 AGAGGAAGCT TCTAGATGAA CTCGGTATCC CACCTGAAGA TGTTCCAGGTT GATAGGTTCA 480 CACACTTTAGG TCGCATGTTG TACAAGGCTC CATCTGATGG AAAATGCGGT GAACATGAAG 540 TTGATTACCT ACTCTTCCTC GTGCGTGACG TTGCCGTGAA CCCAAACCCT GATGAGGTGG 660 CGGACATTAG ATACGTGAAC CAAGAAGAGT TAAAAGGGTT ACTAAGGAAC TTCTTGTTCA 720 AATGGTGGGA TTTGAAATTG TCCCCATGGT TTAGGCTAG CAACACCCT GATGAGCTGG 660 GTGAGGAGGG TTTGAAATTG TCCCCATGGT TAAAAGAGTT ACTAAGGAAC TTCTTGTTCA 720 AATGGTGGGA TCATGCCAA AAGGGGACC TCAATGAAGC AATTGAACCATTC 780 ATAGGTGGA TCATGCCAA AAGGGGACC TCAATGAAGC AATTGAACCATTC 780 ATAAGTTGAT ATGAAAAATG GTTAATATTT ATGGGTGGG | GAAGCTCCTG | GATGAACTCG | GCATCCCTGG | AGCAGATGTT | CCGGTTGATG | AGTTCACTCC | 540 |
| TGTAAAATAT GTGAACCGGG AAGAGCTGAA GGAATTGGTA AGGAAGGCGG ATGCTGGTGA 720 AGAGGGTGTG AAGCTGTCCC CGTGGTTCAA ATTGATTGTC GATAATTTCT TGTTTCAGTG 780 GTGGGATCGA CTCCATAAGG GAACCCTAAC CGAAGCTATT GATATGAAAA CAATCCACAA 840 ACTCACATAA AACACTACA CTAGTAGGA AGAGGATTAT ATGAGATATT TGTTATATGT 900 GAAATTGAAA TTCAGATGAA TGCTTGTATT TATTTCTATT TGGACAAACT TCAACTTCTT 960 TTTGCTACCT TATCAGAAAA AAAAA ***C210 | ATTGGGTCGC | ATTCTATACA | AGGCCGCATC | GGATGGAAAG | TGGGGAGAAC | ATGAACTTGA | 600 |
| AGAGGGTTG AACCTGTCCC CGTGGTTCAA ATTGATTGTC GATAATTTCT TGTTTCAGTG GAGGGGTGGACCG CTCCATAAGG GAACCCTAAC CGAAGCTATT GATATGAAAA CAATCCACAA 840 ACCTCACATAA AACACTACA CTAGTAGGAG AGAGGATTAT ATGAGATATT TGTTATATGT 900 GAAATTGAAA TTCAGATGAA 7GCTTGTATT TATTTCTATT TGGACAAACT TCAACTTCTT 960 TTTGCTACCT TATCAGAAAA AAAAA 9AAAA 985 <211> 988 <212> DNA <211> 988 <212> DNA <213> Lactuca sativa CCATTAACG CCGCCATGGG TGACGATCC GGCATGGACG 120 TTTCTTCAGACAA TCCATCATT TTGATAACT CACCTTCGCC GGTCTTAATC GCTGTCCAGACA ACCTCCATCA TTTCAGACAA TCCATCACC CCGCCATGGG TGACGATTCC GGCATGGACG 120 TTCTTCTGGCA ACCTCCACA TCCACCTTCAGC GGCATGGACG 120 TTCTTCTGGCA ACCTCCACA TCCACCTTCAGC GGCATGGACG 120 TCCTCTCAGA ACGCTCCAATTCT ACCACTTCAGC ACCACTTCAGC ACCATTAACCG CCGCCATGGG TGACGATTCC GGCATGGACG 120 TCCTTCAGAA ACCACTCAT TTGATGATGA AATGCAATT CACTCTCAGC ACCACTCAGA ACCACTATACAA ATACCAATTA CTCCTTCAGC 300 TCCTTCAGAA AACCACTGT TGCAGCAATC 130 AAAGGTCAGA AACCAAGGTG ACAATTCCTT TGGTATGGAA AAACACCTGT TGCAGCCATC 360 AAAGGTCAGA AACCAAGGGT ACATTTCCTT TGGTATGGAA AAACACCTGT TGCAGCCATC 360 AAAAGGTCAGA AACCAAGGGT TCAGAGGAGA TTAATTCCCG AAAATGCCCT TGGGGTCAGA AATGCTGCAC 420 AGAACTATAACCA AAACACCTGT TGCAGCCATC 360 AAACACTTCAGA AAACACCTGT TGCAGCCATC 360 AAACACTTCAGC AAAATGCCCT TGGGGTCAGA AATGCTGCAC 420 AGAACTATAACCA AAACACCTGT TGCAGCCATC CACCTATACCA AAACACCTGT TCCAGAGATCAA CTCCAGACCTT CACCTAGAGA TCCTTCAGC AAAATGCCCT TGGGGTCAGA AATGCTGCAC 420 AAACACCTG AAAATGCGAG TCCAAACCCT AAAAGAGATC AAAAAGACTC AACACCTG AAACACCTG AAAACCATTC AAAGGGGAGA TCATGAAAAAAAACATTC AAAGGGAGAG TCAAACAACACTG AAAACCATTC AAAGAGAGT AAAACACTG AAAACCATTC AAAAGAGGT AAAACACTG AAAACCATTC AAAACACTG AAAACACTTC AA | TTACCTGCTG | TTTATGGTAC | GTGATGTTGG | TTTGGATCCG | AACCCAGATG | AAGTGAAAGA | 660 |
| GTGGGATCGA CTCCATAAGG GAACCCTAAC CGAAGCTATT GATATGAAAA CAATCCACAA ACTCACATAA AAACACTACA CTAGTAGGAG AGAGGATTAT ATGAGATATT TGTTATATGT 900 GAAATTGAAA TTCAGATGAA TGCTTGTATT TATTTCTATT TGGACAAACT TCAACTTCTT 960 TTTGGTACCT TATCAGAAAA AAAAA 985 Columbia</td <td>TGTAAAATAT</td> <td>GTGAACCGGG</td> <td>AAGAGCTGAA</td> <td>GGAATTGGTA</td> <td>AGGAAGGCGG</td> <td>ATGCTGGTGA</td> <td>720</td> | TGTAAAATAT | GTGAACCGGG | AAGAGCTGAA | GGAATTGGTA | AGGAAGGCGG | ATGCTGGTGA | 720 |
| ACTCACATAA AAACACTACA CTAGTAGGAG AGAGGATTAT ATGAGATATT TGTTATATGT 960 GAAATTGAAA TTCAGATGAA TGCTTGTATT TATTTCTATT TGGACAAACT TCAACTTCTT 960 TTTGCTACCT TATCAGAAAA AAAAA 985 Columbia</td <td>AGAGGGTGTG</td> <td>AAGCTGTCCC</td> <td>CGTGGTTCAA</td> <td>ATTGATTGTC</td> <td>GATAATTTCT</td> <td>TGTTTCAGTG</td> <td>780</td> | AGAGGGTGTG | AAGCTGTCCC | CGTGGTTCAA | ATTGATTGTC | GATAATTTCT | TGTTTCAGTG | 780 |
| GAAATTGAAA TTCAGATGAA TGCTTGTATT TATTTCTATT TGGACAAACT TCAACTTCTT 960 TTTGCTACCT TATCAGAAAA AAAAA 9ASAA 985 Columbia</td <td>GTGGGATCGA</td> <td>CTCCATAAGG</td> <td>GAACCCTAAC</td> <td>CGAAGCTATT</td> <td>GATATGAAAA</td> <td>CAATCCACAA</td> <td>840</td> | GTGGGATCGA | CTCCATAAGG | GAACCCTAAC | CGAAGCTATT | GATATGAAAA | CAATCCACAA | 840 |
| CACCATACAGA ACCAAGGTG ACATTTCCTT TAGATACA ACCATTCCAGAGAGA ACCATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA | ACTCACATAA | AAACACTACA | CTAGTAGGAG | AGAGGATTAT | ATGAGATATT | TGTTATATGT | 900 |
| <pre><210> 32 <211> 988 <212> DNA <213> Lactuca sativa </pre> <pre><400> 32 TATTCGCTC AAAATCTCT CCATTAACTG CTCAAATCTC CACCTTCGCC GGTCTTAATC 60 TCCGCCGGCG CACTTTCACC ACCATAACCG CCGCCATGGG TGACGATTCC GGCATGGACG 120 CTGTCCAGAG ACGTCTCATG TTTGATGATG AATGCATTTT GGTTGATGAA AATGACAATG 180 TTCTTGGGCA TGATACCAAA TACAATGTC ACTTGATGAG GAAGATTGAG AAAGATAATT 240 TGCTTCATAG AGCATTCAGT GTATTTTAT TCAATTCAAA ATACGAATTA CTCCTTCAGC 300 AAAGGTCAGA AACCAAGGTG ACATTTCCTT TGGTATGGAC AAACACCTGT TGCAGCCATC 360 CACTATACAG AGAATCGGAG TTAATTCCCG AAAATGCCCT TGGGGTCAGA AATGCTGCAC 420 AGAGGAAGCT TCTAGATGAA CTCGGTATCC CTGCTGAAGA TGTTCCAGTT GATGAGTTCA 480 CAACTTTAGG TCGCATGTG TACAAGGCTC CATCTGATGG AAAATGGGGT GAACATGAAG 540 TTGATTACCT ACTCTTCCTC GTGCGTGACG TTGCCGTGAA CCCAAACCCT GATGAGGTGG 600 CGGACATTAG ATACGTGAAC CAAGAAGAGT TAAAAGAGTT ACTAAGGAAG GCGGATGCGG 660 CTGAGGAGGG TTTGAAATTG TCCCCATGGT TTAGGCTAGT GGTGGACAAC TTCTTGTTCA 720 AATGGTGGGA TCATGTCCAA AAGGGGACA TCAATGAAGC AATTGACATG AAAACCATTC 780 ATAAGGTGGA TCATGTCCAA AAGGGGACAC TCAATGAAGC ATTGACATG AAAACCATTC 780 ATAAGGTGGA TCATGTCCAA AAGGGGACAC TCAATGAAGC ATTGACATG AAAACCATTC 780 ATAAAGTTGAT ATGAAAAATG GTTAATATTT ATGGTGGTGG TTTGCTTTATT TGGGAGGTGTT 990</pre> | GAAATTGAAA | TTCAGATGAA | TGCTTGTATT | TATTTCTATT | TGGACAAACT | TCAACTTCTT | 960 |
| <pre><211> 988 <212> DNA <213> Lactuca sativa </pre> <pre><400> 32 TATTCGCTTC AAAATCTCTT CCATTAACTG CTCAAATCTC CACCTTCGCC GGTCTTAATC 60 TCCGCCGGCG CACTTTCACC ACCATAACCG CCGCCATGGG TGACGATTCC GGCATGGACG 120 CTGTCCAGAG ACGTCCATG TTTGATGATG AATGCATTTT GGTTGATGAA AATGACAATG 180 TTCTTGGGCA TGATACCAAA TACAATTGTC ACTTGATGGA GAAGATTGAG AAAGATAATT 240 TGCTTCATAG AGCATCAGT GTATTTTAT TCAATTCAAA ATACGAATTA CTCCTTCAGC 300 AAAGGTCAGA AACCAAGGTG ACATTTCCTT TGGTATGGAC AAACACCTGT TGCAGCCATC 360 CACTATACAG AGAATCGGAG TTAATTCCCG AAAATGCCCT TGGGGTCAGA AATGCTGCAC 420 AGAGGAAGCT TCTAGATGAA CTCGGTATCC CTGCTGAAGA TGTTCCAGTT GATGAGTTCA 480 CAACTTTAGG TCGCATGTTG TACAAGGCTC CATCTGATGG AAAATGGGGT GAACATGAAG 540 TTGATTACCT ACTCTTCCTC GTGCGTGACG TTGCCGTGAA CCCCAAACCCT GATGAGGTGG 600 CGGACATTAG ATACGGAAC CAAGAAGAGT TAAAAGAGTT ACTAAGGAAG GCGGATGCGG 660 GTGAGGAGGG TTTGAAATTG TCCCCATGGT TAAGAGAGT ACTAAGGAAC TTCTTGTTCA 720 AATGGTGGGA TCATGTCCAA AAGGGGACAC TCAATGAAGC AATTGACAT AAAACCATTC 780 ATAAGTTGAT ATGAAAAATG GTTAATATTT ATGGTGGTGG TTTGGAGCTTA ATAATTTGTG 840 TGTTCAAGTC TCGGTCCTC TTTTTTTAAC GTTTTTTTT TTTCTTTTAT TGGGAGGTGTT 900</pre> | TTTGCTACCT | TATCAGAAAA | AAAAA | | | | 985 |
| TATTCGCTTC AAAATCTCTT CCATTAACTG CTCAAATCTC CACCTTCGCC GGTCTTAATC 60 TCCGCCGGCG CACTTTCACC ACCATAACCG CCGCCATGGG TGACGATTCC GGCATGGACG 120 CTGTCCAGAG ACGTCTCATG TTTGATGATG AATGCATTTT GGTTGATGAA AATGACAATG 180 TTCTTGGGCA TGATACCAAA TACAATTGTC ACTTGATGGA GAAGATTGAG AAAGATAATT 240 TGCTTCATAG AGCATTCAGT GTATTTTAT TCAATTCAAA ATACGAATTA CTCCTTCAGC 300 AAAGGTCAGA AACCAAGGTG ACATTCCCT TGGTATGGAC AAACACCTGT TGCAGCCATC 360 CACTATACAG AGAATCGGAG TTAATTCCCG AAAATGCCCT TGGGGTCAGA AATGCTGCAC 420 AGAGGAAGCT TCTAGATGAA CTCGGTATCC CTGCTGAAGA TGTTCCAGTT GATGAGTTCA 480 CAACTTTAGG TCGCATGTTG TACAAGGCTC CATCTGATGG AAAATGGGGT GAACATGAAG 540 CTGGACATTAG ATACGTGAAC CAAGAAGAGT TAAAAGAGTT ACTAAGGAAG GCGGATGCGG 660 CGGACATTAG ATACGTGAAC CAAGAAGAGT TAAAAGAGTT ACTAAGGAAG GCGGATGCGG 660 CTGAGGGAGGG TTTGAAATTG TCCCCATGGT TTAGGCTAGT GGTGGACAAC TTCTTGTTCA 720 AATGGTGGGA TCATGTCCAA AAGGGGACAC TCAATGAAGC AATTGACATG AAAACCATTC 780 ATAAGTTGAT ATGAAAAATG GTTAATATTT ATGGTGGTGG TTTGGAGCTA ATAATTTGTG 840 TGTTCAAGTC TCGGTCCTTC TTTTTTAAC GTTTTTTTT TTTCTTTTAT TGGGAGTGTT 900 | <211> 988 <212> DNA | uca sativa | | | | | |
| TCCGCCGGCG CACTTCACC ACCATAACCG CCGCCATGGG TGACGATTCC GGCATGGACG 120 CTGTCCAGAG ACGTCTCATG TTTGATGATG AATGCATTTT GGTTGATGAA AATGCAATG 180 TTCTTGGGCA TGATACCAAA TACAATTGTC ACTTGATGGA GAAGATTGAG AAAGATAATT 240 TGCTTCATAG AGCATTCAGT GTATTTTAT TCAATTCAAA ATACGAATTA CTCCTTCAGC 300 AAAAGGTCAGA AACCAAGGTG ACATTCCTT TGGTATGGAC AAACACCTGT TGCAGCCATC 360 CACTATACAG AGAATCGGAG TTAATTCCCG AAAATGCCCT TGGGGTCAGA AATGCTGCAC 420 AGAGGAAGCT TCTAGATGAA CTCGGTATCC CTGCTGAAGA TGTTCCAGTT GATGAGTTCA 480 CAACTTTAGG TCGCATGTTG TACAAGGCTC CATCTGATGG AAAATGGGGT GAACATGAAG 540 TTGATTACCT ACTCTCCCT GTGCGTGACG TTGCCGTGAA CCCCAAACCCT GATGAGGTGG 600 CGGACATTAG ATACGTGAAC CAAGAAGAGT TAAAAAGAGTT ACTAAGGAAG GCGGATGCGG 660 GTGAGGAGGG TTTGAAAATTG TCCCCATGGT TTAGGCTAGT GGTGGACAAC TTCTTGTTCA 720 AATGGTGGGA TCATGTCCAA AAGGGGACAC TCAATGAAGC AATTGACATG AAAACCATTC 780 ATAAGTTGAT ATGAAAAATG GTTAATATTT ATGGTGGTGG TTTGGAGCTA ATAATTTGTG 840 TGTTCAAGTC TCGGTCCTTC TTTTTTTAAC GTTTTTTTT TTTCTTTTAT TGGGAGTGTT 900 | | 7 7 7 7 MCMCMM | | CMC2 2 2 mone | | | |
| CTGTCCAGAG ACGTCTCATG TTTGATGATG AATGCATTT GGTTGATGAA AATGACAATG 180 TTCTTGGGCA TGATACCAAA TACAATTGTC ACTTGATGGA GAAGATTGAG AAAGATAATT 240 TGCTTCATAG AGCATTCAGT GTATTTTAT TCAATTCAAA ATACGAATTA CTCCTTCAGC 300 AAAGGTCAGA AACCAAGGTG ACATTTCCTT TGGTATGGAC AAACACCTGT TGCAGCCATC 360 CACTATACAG AGAATCGGAG TTAATTCCCG AAAATGCCCT TGGGGTCAGA AATGCTGCAC 420 AGAGGAAGCT TCTAGATGAA CTCGGTATCC CTGCTGAAGA TGTTCCAGTT GATGAGTTCA 480 CAACTTTAGG TCGCATGTTG TACAAGGCTC CATCTGATGG AAAATGGGGT GAACATGAAG 540 TTGATTACCT ACTCTTCCTC GTGCGTGACG TTGCCGTGAA CCCCAAACCCT GATGAGGTGG 600 CGGACATTAG ATACGTGAAC CAAGAAGAGT TAAAAGAGTT ACTAAGGAAG GCGGATGCGG 660 GTGAGGAGGG TTTGAAATTG TCCCCATGGT TTAGGCTAGT GGTGGACAAC TTCTTGTTCA 720 AATGGTGGGA TCATGTCCAA AAGGGGACAC TCAATGAAGC AATTGACATG AAAACCATTC 780 ATAAGTTGAT ATGAAAAATG GTTAATATTT ATGGTGGTGG TTTGGAGCTA ATAATTTTGTG 840 TGTTCAAGTC TCGGTCCTC TTTTTTTAC GTTTTTTTTT TTTCTTTTAT TGGGAGTGTT 900 | | | | | | | 60 |
| TTCTTGGCA TGATACCAAA TACAATTGTC ACTTGATGA GAAGATTGAG AAAGATAATT 240 TGCTTCATAG AGCATTCAGT GTATTTTAT TCAATTCAAA ATACGAATTA CTCCTTCAGC 300 AAAGGTCAGA AACCAAGGTG ACATTTCCTT TGGTATGGAC AAACACCTGT TGCAGCCATC 360 CACTATACAG AGAATCGGAG TTAATTCCCG AAAATGCCCT TGGGGTCAGA AATGCTGCAC 420 AGAGGAAGCT TCTAGATGAA CTCGGTATCC CTGCTGAAGA TGTTCCAGTT GATGAGTTCA 480 CAACTTTAGG TCGCATGTTG TACAAGGCTC CATCTGATGG AAAATGGGGT GAACATGAAG 540 TTGATTACCT ACTCTTCCTC GTGCGTGACG TTGCCGTGAA CCCCAAACCCT GATGAGGTGG 660 CGGACATTAG ATACGTGAAC CAAGAAGAGT TAAAAGAGTT ACTAAGGAAG GCGGATGCGG 660 GTGAGGAGGG TTTGAAATTG TCCCCATGGT TTAGGCTAGT GGTGGACAAC TTCTTGTTCA 720 AATGGTGGGA TCATGTCCAA AAGGGGACAC TCAATGAAGC AATTGACATG AAAACCATTC 780 ATAAGTTGAT ATGAAAAATG GTTAATATTT ATGGTGGTGG TTTGGAGCTA ATAATTTGTG 840 TGTTCAAGTC TCGGTCCTTC TTTTTTAAC GTTTTTTTTT TTTCTTTTAT TGGGAGTGTT 900 | | | | | | | |
| TGCTTCATAG AGCATTCAGT GTATTTTAT TCAATTCAAA ATACGAATTA CTCCTTCAGC 300 AAAGGTCAGA AACCAAGGTG ACATTTCCTT TGGTATGGAC AAACACCTGT TGCAGCCATC 360 CACTATACAG AGAATCGGAG TTAATTCCCG AAAATGCCCT TGGGGTCAGA AATGCTGCAC 420 AGAGGAAGCT TCTAGATGAA CTCGGTATCC CTGCTGAAGA TGTTCCAGTT GATGAGTTCA 480 CAACTTTAGG TCGCATGTTG TACAAGGCTC CATCTGATGG AAAATGGGGT GAACATGAAG 540 TTGATTACCT ACTCTTCCTC GTGCGTGACG TTGCCGTGAA CCCCAAACCCT GATGAGGTGG 600 CGGACATTAG ATACGTGAAC CAAGAAGAGT TAAAAGAGTT ACTAAGGAAG GCGGATGCGG 660 GTGAGGAGGG TTTGAAATTG TCCCCATGGT TTAGGCTAGT GGTGGACAAC TTCTTGTTCA 720 AATGGTGGGA TCATGTCCAA AAGGGGACAC TCAATGAAGC AATTGACATG AAAACCATTC 780 ATAAGTTGAT ATGAAAAAATG GTTAATATTT ATGGTGGTGG TTTTGGAGCTA ATAATTTTGTG 840 TGTTCAAGTC TCGGTCCTTC TTTTTTAAC GTTTTTTTTT TTTCTTTTAT TGGGAGTGTT 900 | | | | | | | 180 |
| AAAGGTCAGA AACCAAGGTG ACATTTCCTT TGGTATGGAC AAACACCTGT TGCAGCCATC 360 CACTATACAG AGAATCGGAG TTAATTCCCG AAAATGCCCT TGGGGTCAGA AATGCTGCAC 420 AGAGGAAGCT TCTAGATGAA CTCGGTATCC CTGCTGAAGA TGTTCCAGTT GATGAGTTCA 480 CAACTTTAGG TCGCATGTTG TACAAGGCTC CATCTGATGG AAAATGGGGT GAACATGAAG 540 TTGATTACCT ACTCTTCCTC GTGCGTGACG TTGCCGTGAA CCCCAAACCCT GATGAGGTGG 600 CGGACATTAG ATACGTGAAC CAAGAAGAGT TAAAAGAGTT ACTAAGGAAG GCGGATGCGG 660 GTGAGGAGGG TTTGAAATTG TCCCCATGGT TTAGGCTAGT GGTGGACAAC TTCTTGTTCA 720 AATGGTGGGA TCATGTCCAA AAGGGGACAC TCAATGAAGC AATTGACATG AAAACCATTC 780 ATAAGTTGAT ATGAAAAATG GTTAATATTT ATGGTGGTGG TTTGGAGCTA ATAATTTGTG 840 TGTTCAAGTC TCGGTCCTC TTTTTTAAC GTTTTTTTT TTTCTTTTAT TGGGAGTGTT 900 | | | | | | • | 240 |
| CACTATACAG AGAATCGGAG TTAATTCCCG AAAATGCCCT TGGGGTCAGA AATGCTGCAC 420 AGAGGAAGCT TCTAGATGAA CTCGGTATCC CTGCTGAAGA TGTTCCAGTT GATGAGTTCA 480 CAACTTTAGG TCGCATGTTG TACAAGGCTC CATCTGATGG AAAATGGGGT GAACATGAAG 540 TTGATTACCT ACTCTTCCTC GTGCGTGACG TTGCCGTGAA CCCCT GATGAGGTGG 600 CGGACATTAG ATACGTGAAC CAAGAAGAGT TAAAAGAGTT ACTAAGGAAG GCGGATGCGG 660 GTGAGGAGGG TTTGAAATTG TCCCCATGGT TTAGGCTAGT GGTGGACAAC TTCTTGTTCA 720 AATGGTGGGA TCATGTCCAA AAGGGGACAC TCAATGAAGC AATTGACATG AAAACCATTC 780 ATAAGTTGAT ATGAAAATG GTTAATATTT ATGGTGGTGG TTTGGAGCTA ATAATTTGTG 840 TGTTCAAGTC TCGGTCCTTC TTTTTTAAC GTTTTTTTTT TTTCTTTTAT TGGGAGTGTT 900 | | | | | | | 300 |
| AGAGGAAGCT TCTAGATGAA CTCGGTATCC CTGCTGAAGA TGTTCCAGTT GATGAGTTCA 480 CAACTTTAGG TCGCATGTTG TACAAGGCTC CATCTGATGG AAAATGGGGT GAACATGAAG 540 TTGATTACCT ACTCTTCCTC GTGCGTGACG TTGCCGTGAA CCCCAAACCCT GATGAGGTGG 600 CGGACATTAG ATACGTGAAC CAAGAAGAGT TAAAAGAGTT ACTAAGGAAG GCGGATGCGG 660 GTGAGGAGGG TTTGAAATTG TCCCCATGGT TTAGGCTAGT GGTGGACAAC TTCTTGTTCA 720 AATGGTGGGA TCATGTCCAA AAGGGGACAC TCAATGAAGC AATTGACATG AAAACCATTC 780 ATAAGTTGAT ATGAAAAATG GTTAATATTT ATGGTGGTGG TTTGGAGCTA ATAATTTGTG 840 TGTTCAAGTC TCGGTCCTTC TTTTTTAAC GTTTTTTTTT TTTCTTTTAT TGGGAGTGTT 900 | | | | | | | 360 |
| CAACTTTAGG TCGCATGTTG TACAAGGCTC CATCTGATGG AAAATGGGGT GAACATGAAG 540 TTGATTACCT ACTCTCCTC GTGCGTGACG TTGCCGTGAA CCCCAAACCCT GATGAGGTGG 600 CGGACATTAG ATACGTGAAC CAAGAAGAGT TAAAAGAGTT ACTAAGGAAG GCGGATGCGG 660 GTGAGGAGGG TTTGAAATTG TCCCCATGGT TTAGGCTAGT GGTGGACAAC TTCTTGTTCA 720 AATGGTGGGA TCATGTCCAA AAGGGGACAC TCAATGAAGC AATTGACATG AAAACCATTC 780 ATAAGTTGAT ATGAAAAATG GTTAATATTT ATGGTGGTGG TTTGGAGCTA ATAATTTGTG 840 TGTTCAAGTC TCGGTCCTTC TTTTTTAAC GTTTTTTTTT TTGTGTTTTAT TGGGAGTGTT 900 | | | | | | | 420 |
| TTGATTACCT ACTCTCCTC GTGCGTGACG TTGCCGTGAA CCCCAAACCCT GATGAGGTGG 600 CGGACATTAG ATACGTGAAC CAAGAAGAGT TAAAAGAGTT ACTAAGGAAG GCGGATGCGG 660 GTGAGGAGGG TTTGAAATTG TCCCCATGGT TTAGGCTAGT GGTGGACAAC TTCTTGTTCA 720 AATGGTGGGA TCATGTCCAA AAGGGGACAC TCAATGAAGC AATTGACATG AAAACCATTC 780 ATAAGTTGAT ATGAAAATG GTTAATATTT ATGGTGGTGG TTTGGAGCTA ATAATTTGTG 840 TGTTCAAGTC TCGGTCCTTC TTTTTTAAC GTTTTTTTT TTTCTTTTAT TGGGAGTGTT 900 | | | | | | | 480 |
| CGGACATTAG ATACGTGAAC CAAGAAGAGT TAAAAGAGTT ACTAAGGAAG GCGGATGCGG 660 GTGAGGAGGG TTTGAAATTG TCCCCATGGT TTAGGCTAGT GGTGGACAAC TTCTTGTTCA 720 AATGGTGGGA TCATGTCCAA AAGGGGACAC TCAATGAAGC AATTGACATG AAAACCATTC 780 ATAAGTTGAT ATGAAAAATG GTTAATATTT ATGGTGGTGG TTTGGAGCTA ATAATTTGTG 840 TGTTCAAGTC TCGGTCCTTC TTTTTTTAAC GTTTTTTTT TTTCTTTTAT TGGGAGTGTT 900 | | | | | | | 540 |
| GTGAGGAGGG TTTGAAATTG TCCCCATGGT TTAGGCTAGT GGTGGACAAC TTCTTGTTCA 720 AATGGTGGGA TCATGTCCAA AAGGGGACAC TCAATGAAGC AATTGACATG AAAACCATTC 780 ATAAGTTGAT ATGAAAAATG GTTAATATTT ATGGTGGTGG TTTGGAGCTA ATAATTTGTG 840 TGTTCAAGTC TCGGTCCTTC TTTTTTAAC GTTTTTTTT TTTCTTTTAT TGGGAGTGTT 900 | TTGATTACCT | ACTCTTCCTC | GTGCGTGACG | TTGCCGTGAA | CCCAAACCCT | GATGAGGTGG | 600 |
| AATGGTGGGA TCATGTCCAA AAGGGGACAC TCAATGAAGC AATTGACATG AAAACCATTC 780 ATAAGTTGAT ATGAAAAATG GTTAATATTT ATGGTGGTGG TTTGGAGCTA ATAATTTGTG 840 TGTTCAAGTC TCGGTCCTTC TTTTTTAAC GTTTTTTTT TTTCTTTTAT TGGGAGTGTT 900 | CGGACATTAG | ATACGTGAAC | CAAGAAGAGT | TAAAAGAGTT | ACTAAGGAAG | GCGGATGCGG | 660 |
| ATAAGTTGAT ATGAAAAATG GTTAATATTT ATGGTGGTGG TTTGGAGCTA ATAATTTGTG 840 TGTTCAAGTC TCGGTCCTTC TTTTTTTAAC GTTTTTTTT TTTCTTTTAT TGGGAGTGTT 900 | GTGAGGAGGG | TTTGAAATTG | TCCCCATGGT | TTAGGCTAGT | GGTGGACAAC | TTCTTGTTCA | 720 |
| TGTTCAAGTC TCGGTCCTTC TTTTTTTAAC GTTTTTTTT TTTCTTTTAT TGGGAGTGTT 900 | AATGGTGGGA | TCATGTCCAA | AAGGGGACAC | TCAATGAAGC | AATTGACATG | AAAACCATTC | 780 |
| | ATAAGTTGAT | ATGAAAAATG | GTTAATATTT | ATGGTGGTGG | TTTGGAGCTA | ATAATTTGTG | 840 |
| TATTGTGTAC TTGTAACGTA GGCCCTTTGG TTACGCTTTA AGAGTTTAAT AAAGAACCAC 960 | TGTTCAAGTC | TCGGTCCTTC | TTTTTTTAAC | GTTTTTTTT | TTTCTTTTAT | TGGGAGTGTT | 900 |
| | TATTGTGTAC | TTGTAACGTA | GGCCCTTTGG | TTACGCTTTA | AGAGTTTAAT | AAAGAACCAC | 960 |

PCT/US99/12121

840

900

960

1020

1080

1140

1200

1260

1320

1380

1440

1500

1560

1620



ССТТААТТТА ААААААААА ААААААА 988 <210> 33 <211> 1874 <212> DNA <213> Chlamydomonas reinhardtii GGCACGAGCT CGAGTTTGTT TTACCATGAC ATCGGGAATT TGGAAGCTTG AACTACCTCA 60 ATTACTCAAG TAACTCGCGG CAACACATTT CGCGCGCCAT CGCTGTTTTC TCTGCTCCAG 120 CTACCGAGCA GCATTGCTTT AGATCGCTTT GATGTCATAA ACTCCCACTT ATATGAGATC 180 CAGTTTCATC GAGCCCAAGC CCAGAGCGCA ACCTGTCTTA AGCCGCGGCA GGGCGTCCAT 240 GCGCCTCGCG CAAAGCCGTG CTCTCGTTGC GCGTGTCAGC TCCGCCCTGT GGCCGGGAGC 300 AGGACTTTCA CAGGCTCAAA GCGTTGCGGT GCGAATGGCG AGTTCGTCAA CCTGGGAAGG 360 CACGGGCCTG AGCCAGGATG ACTTCATGCA GCGGGACGAG TGCTTGGTGG TGGACGAGCA 420 GGACCGGCTG CTAGGCACCG CCAACAAGTA CGACTGCCAC CGCTTCGAGG CGGCCAAGGG 480 CCAGCCCTGC GGCCGCCTGC ACCGCGCCTT CTCCGTGTTC CTGTTCAGCC CCGACGGCCG 540 ACTGCTGCTG CAGCAGCGC CAGCCAGCAA GGTGACGTTC CCGGGTGTGT GGACCAACAC 600 CTGCTGCTCG CACCCGCTGG CGGGCCAGGC GCCGGACGAG GTGGACCTGC CGGCGGCGGT 660 AGCCTCGGGC CAGGTGCCGG GCATCAAGGC GGCGGCGGTG CGCAAGCTGC AGCACGAGCT 720 GGGGATACCG CCGGAGCAGG TTCCCGCCTC CTCCTTCTCC TTCCTCACGC GTCTGCACTA 780

CTGCGCCGCC GACACCGCCA CGCACGGCCC GGCGGCGGAG TGGGGCGAGC ACGAGGTGGA

CTACGTGCTG TTCGTGCGGC CGCAGCAGCC CGTCAGCCTG CAGCCCAACC CAGACGAGGT

GGACGCCACG CGCTACGTGA CGCTGCCGGA GCTTCAGTCC ATGATGGCGG ACCCCGGCCT

CAGCTGGAGC CCCTGGTTCC GCATCCTGGC CACACAGCCC GCCTTCCTGC CCGCCTGGTG

GGGCGACCTG AAGCGGCGCT GGCGCCCGGG CGGCAGCCGA CTGTCGGACT GGGGCACCAT

CCACCGCGTC ATGTGAAGAA AAAGGGGAAG CAGGGGGGGG AGCGGGGGAT GAATGGGAAT

GTGAATGCGA TTGTGATGCG GCGTGGGATG AGGTCTGAAG ACAGGGGGGAA AATCGGGGGG

CGGGCGTGAG CGTGTGTGTA CGTGAGCGAC AAAGCCGGGA GGCGGACCGC GCGATGGGTA

CATGTGTGTG CGGAGGGTCG GTGGGTCGGT CGGTTGCGCG GCATAGCGTG TTGTGTGTGT

GCGGCTGCAG GGGTATGTGG GCACCCGGGC ACGGAGGAGA AGGCACACGC AGGTGGCGCG

GAGGTGTGTC AGGGGCCCATG GGCGGGCCTC ACTCCTGGTC GTGCCCAGTG GTCTCGTGGG

CAGAGTGGCA GGGGCTGCAC CCATATGAGC GGCGCACTGC CGCGCTGGGC TAAGTCCTTA

TCACTTGGTG AGGTGGGCG AGGTGGCTGT GGGCGGCGGG CGCAGTGGCA GAAGGACACG

GTGTGTGAGC GGTGGAGCTC TGGCCGTGCC GGCCGTGAGG GGCGGATAGC GATATGACGT

| 1110 00/600 | | | D.OTT. (1000 (1010) |
|-------------|---|-------|---------------------|
| WO 99/63055 | _ | , | PCT/US99/12121 |

| TGTGCTTGGC | CGCTGTAATG | CGGGAGAATG | TGCAGGCCGC | GAGAAGCGGG | CGGTGGCAGG | 1680 |
|--|------------|------------|------------|------------|------------|------|
| AGGCCGCAGG | CTGCAGCACC | CGTTĠGGGAG | GTGCCACCTG | CAGGCGCGGC | GCCGGGCGGG | 1740 |
| CCTGAGTAAT | GGGCGCCTGA | GTAGTGGCGG | CCACAGGAGG | CGCAGGAGGC | AGCAGCAGGA | 1800 |
| GGACGAGCTG | GAGGGACCCG | TTGGCAACCC | AAGGTTGCGC | GTGTAACATA | GTGGCCATAC | 1860 |
| AAAAAAAA | AAAA | | | | | 1874 |
| <210> 34 <211> 954 <212> DNA <213> Taget | es erecta | | | | · | |
| <400> 34 CCAAAAACAA | СТСАААТСТС | CTCCGTCGCT | CTTACTCCGC | CATGGGTGAC | GACTCCGGCA | 60 |
| TGGATGCTGT | TCAGCGACGT | CTCATGTTTG | ACGATGAATG | CATTTTGGTG | GATGAGTGTG | 120 |
| ACAATGTGGT | GGGACATGAT | ACCAAATACA | ATTGTCACTT | GATGGAGAAG | ATTGAAACAG | 180 |
| GTAAAATGCT | GCACAGAGCA | TTCAGCGTTT | TTCTATTCAA | TTCAAAATAC | GAGTTACTTC | 240 |
| TTCAGCAACG | GTCTGCAACC | AAGGTGACAT | TTCCTTTAGT | ATGGACCAAC | ACCTGTTGCA | 300 |
| GCCATCCACT | CTACAGAGAA | TCCGAGCTTG | TTCCCGAAAA | CGCCCTTGGA | GTAAGAAATG | 360 |
| CTGCACAGAG | GAAGCTGTTG | GATGAACTCG | GTATCCCTGC | TGAAGATGTT | CCCGTTGATC | 420 |
| AGTTTACTCC | TTTAGGTCGC | ATGCTCTACA | AGGCTCCATC | TGATGGAAAG | TGGGGAGAAC | 480 |
| ATGAACTTGA | СТАССТАСТТ | TTCATAGTGA | GAGACGTTGC | TGTAAACCCG | AACCCAGATG | 540 |
| AAGTGGCGGA | TATCAAATAT | GTGACCAGAA | GAGTTAAAGG | AGCTGCTAAG | GAAAGCAGAT | 600 |
| GCGGGGGAGG | AGGGTTTGAA | GCTGTCTCCA | TGGTTCAGGT | TAGTGGTTGA | TAACTTCTTG | 660 |
| TTCAAGTGGT | GGGATCATGT | GCAAAAGGGT | ACACTCACTG | AAGCAATTGA | TATGAAAACC | 720 |
| ATACACAAGC | TGATATAGAA | ACACACCCTC | AACCGAAAAG | TTCAAGCCTA | ATAATTCGGG | 780 |
| TTGGGTCGGG | TCTACCATCA | ATTGTTTTTT | TCTTTTAAGA | AGTTTTAATC | TCTATTTGAG | 840 |
| CATGTTGATT | CTTGTCTTTT | GTGTGTAAGA | TTTTGGGTTT | CGTTTCAGTT | GTAATAATGA | 900 |
| ACCATTGATG | GTTTGCAATT | TCAAGTTCCT | ATCGACATGT | AGTGATCTAA | AAAA | 954 |
| <210> 35 <211> 1031 <212> DNA <213> Oryza | a sativa | | | | | |
| <400> 35 CCTCCCTTTG | CCTCGCGCAG | AGGCGGCCGC | GCCTTCTCCG | CCGCGAGGAT | GGCCGGCGCC | 60 |
| GCCGCCGCCG | TGGAGGACGC | CGGGATGGAC | GAGGTCCAGA | AGCGGCTCAT | GTTCGACGAC | 120 |
| GAATGCATTT | TGGTGGATGA | ACAAGACAAT | GTTGTTGGCC | ATGAATCAAA | ATATAACTGC | 180 |
| CATCTGATGG | АААААТСБА | ATCTGAAAAT | CTACTTCATA | GGGCTTTCAG | TGTATTCCTG | 240 |

WO 99/63055



PCT/US99/12121

| TTCAACTCAA | AATATGAACT | CCTACTCCAG | CAACGATCTG | CAACAAAGGT | TACATTTCCT | 300 |
|------------|------------|------------|------------|------------|------------|------|
| CTAGTTTGGA | CCAACACTTG | CTGCAGCCAT | CCTCTGTACC | GTGAGTCTGA | GCTTATACAG | 360 |
| GAAAACTACC | TTGGTGTTAG | AAATGCTGCT | CAGAGGAAGC | TCTTGGATGA | GCTGGGCATC | 420 |
| CCAGCTGAAG | ATGTGCCAGT | TGACCAATTC | ACCCCTCTTG | GTCGGATGCT | TTACAAGGCC | 480 |
| CCATCTGATG | GAAAATGGGG | TGAACACGAG | CTTGACTACC | TGCTGTTCAT | CGTCCGCGAC | 540 |
| GTGAAGGTAG | TCCCGAACCC | GGACGAAGTG | GCCGATGTGA | AATACGTGAG | CCGTGAGCAG | 600 |
| CTGAAGGAGC | TCATCCGCAA | AGCGGACGCC | GGAGAGGAAG | GCCTGAAGCT | GTCTCCCTGG | 660 |
| TTCCGGCTGG | TTGTTGACAA | CTTCCTCATG | GGCTGGTGGG | ATCACGTCGA | GAAAGGCACC | 720 |
| CTCAACGAGG | CCGTGGACAT | GGAGACCATC | CACAAGCTGA | AGTAAGGACT | GCGATGTTGT | 780 |
| GGCTGGAAAG | AATGATCCTG | AAGACTCTGT | тсттстсстс | CTGCATATTA | CTCTTACCAG | 840 |
| GGAAGTTGCA | GAAGTCAGAA | GAAGCTTTTG | TATGTTTCTG | GGTTTGGAGC | TTGGAAGTGT | 900 |
| TGGGCTCTGC | TGACTGAGAG | ATTCCCTTAT | AGAGTGTCTA | TGTTAATTTA | GCAAACTTCT | 960 |
| ATATTATACA | TGATTAGTTA | ATTGTTCGGT | GTCTGAATAA | AGAACAATAG | CATGTTCCAT | 1020 |
| GTTTATTTGC | T | | | | | 1031 |
| | | | | • | | |

<210> 36 <211> 232

<212> PRT

<213> Tagetes erecta

<400> 36

Met Gly Asp Asp Ser Gly Met Asp Ala Val Gln Arg Arg Leu Met Phe

Asp Asp Glu Cys Ile Leu Val Asp Glu Cys Asp Asn Val Val Gly His,

Asp Thr Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Thr Gly Lys

Met Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu

Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro Leu Val

Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser Glu Leu 85 90

Val Pro Glu Asn Ala Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu

Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp Gln Phe 115

Thr Pro Leu Gly Arg Met Leu Tyr Lys Ala Pro Ser Asp Gly Lys Trp 135



- Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Ile Val Arg Asp Val Ala 145 150 155 160
- Val Asn Pro Asn Pro Asp Glu Val Ala Asp Ile Lys Tyr Val Ser His 165 170 175
- Glu Glu Leu Lys Glu Leu Leu Arg Lys Ala Asp Ala Gly Glu Glu Gly 180 185 190
- Leu Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe Leu Phe 195 200 205
- Lys Trp Trp Asp His Val Gln Lys Gly Thr Leu Thr Glu Ala Ile Asp 210 215 220
- Met Lys Thr Ile His Lys Leu Ile 225 230
- <210> 37
- <211> 280
- <212> PRT
- <213> Lactuca Sativa
- <400> 37
- Met Leu Lys Phe Pro Pro Phe Lys Thr Ile Ala Thr Met Ile Ser Ser 1 5 10 15
- Pro Tyr Ser Ser Phe Leu Leu Pro Arg Lys Ser Ser Phe Pro Pro Met 20 25 30
- Pro Ser Leu Ala Ala Ala Ser Val Phe Leu His Pro Leu Ser Ser Ala 35 40 45
- Ala Met Gly Asp Ser Ser Met Asp Ala Val Gln Arg Arg Leu Met Phe 50 55 60
- Asp Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Lys Val Val Gly His 65 70 75 80
- Asp Thr Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Lys Gly Asn 85 90 95
- Met Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu 100 105 110
- Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro Leu Val 115 120 125
- Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser Glu Leu 130 135 140
- Ile Asp Glu Asn Ala Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu 145 150 155 160
- Leu Asp Glu Leu Gly Ile Pro Gly Ala Asp Val Pro Val Asp Glu Phe
 165 170 175
- Thr Pro Leu Gly Arg Ile Leu Tyr Lys Ala Ala Ser Asp Gly Lys Trp 180 185 190
- Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Met Val Arg Asp Val Gly

F 1

195

9/63055

200 205

Leu Asp Pro Asn Pro Asp Glu Val Lys Asp Val Lys Tyr Val Asn Arg 210 215 220

Glu Glu Leu Lys Glu Leu Val Arg Lys Ala Asp Ala Gly Glu Gly 225 230 235 240

Val Lys Leu Ser Pro Trp Phe Lys Leu Ile Val Asp Asn Phe Leu Phe 245 250 255

Gln Trp Trp Asp Arg Leu His Lys Gly Thr Leu Thr Glu Ala Ile Asp 260 265 270

Met Lys Thr Ile His Lys Leu Thr 275 280

<210> 38

<211> 229

<212> PRT

<213> Lactuca Sativa

<400> 38

Met Gly Asp Asp Ser Gly Met Asp Ala Val Gln Arg Arg Leu Met Phe 1 5 10 15

Asp Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Asn Val Leu Gly His 20 25 30

Asp Thr Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Lys Asp Asn 35 40 45

Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu 50 55 60

Leu Leu Leu Gln Gln Arg Ser Glu Thr Lys Val Thr Phe Pro Leu Val 65 70 75 80

Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser Glu Leu 85 90 95

Ile Pro Glu Asn Ala Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu 100 105 110

Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp Glu Phe 115 120 125

Thr Thr Leu Gly Arg Met Leu Tyr Lys Ala Pro Ser Asp Gly Lys Trp 130 135 140

Gly Glu His Glu Val Asp Tyr Leu Leu Phe Leu Val Arg Asp Val Ala 145 150 155 160

Val Asn Pro Asn Pro Asp Glu Val Ala Asp Ile Arg Tyr Val Asn Gln 165 170 175

Glu Glu Leu Lys Glu Leu Leu Arg Lys Ala Asp Ala Gly Glu Gly 180 185 190

Leu Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe Leu Phe 195 200 205

Lys Trp Trp Asp His Val Gln Lys Gly Thr Leu Asn Glu Ala Ile Asp 210 225 220

Met Lys Thr Ile His

<210> 39

<211> 295

<212> PRT

<213> Adonis Palaestina

<400> 39

Met Ser Ser Ile Arg Ile Asn Pro Leu Tyr Ser Ile Phe Ser Thr Thr
1 5 10 15

Thr Lys Thr Leu Ser Ala Ser Cys Ser Ser Pro Ala Val His Leu Gln 20 25 30

Gln Arg Cys Arg Thr Leu Ser Ile Ser Ser Ser Ile Thr Asn Ser Pro 35 40 45

Arg Arg Gly Leu Asn Arg Leu Phe Ala Ser Thr Ser Thr Met Gly Glu 50 60

Val Ala Asp Ala Gly Met Asp Ala Val Gln Lys Arg Leu Met Phe Asp 65 70 75 80

Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Lys Val Val Gly Tyr Asp 85 90 95

Ser Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala Glu Asn Leu 100 105 110

Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu Leu 115 120 125

Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro Leu Val Trp 130 135 140

Thr Asn Thr Cys Cys Ser His Pro Leu Phe Arg Asp Ser Glu Leu Ile 145 150 155 160

Glu Glu Asn Phe Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu Leu 165 170 175

Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp Glu Phe Thr 180 185 190

Pro Leu Gly Arg Ile Leu Tyr Lys Ala Pro Ser Asp Gly Lys Trp Gly 195 200 205

Glu His Glu Leu Asp Tyr Leu Leu Phe Ile Val Arg Asp Val Lys Tyr 210 215 220

Asp Pro Asn Pro Asp Glu Val Ala Asp Ala Lys Tyr Val Asn Arg Glu 225 230 235 240

Glu Leu Lys Glu Ile Leu Arg Lys Ala Asp Ala Gly Glu Glu Gly Ile
245 250 255

Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe Leu Phe Lys

265

270

Trp Trp Asp His Val Glu Glu Gly Lys Ile Lys Asp Val Ala Asp Met 275 280 285

Lys Thr lle His Lys Leu Thr 290 295

<210> 40

<211> 234

<212> PRT

<213> Adonis Palaestina

<400> 40

Met Gly Glu Val Thr Asp Ala Gly Met Asp Ala Val Gln Lys Arg Leu $1 \hspace{1cm} 5 \hspace{1cm} 10 \hspace{1cm} 15$

Met Phe Asp Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Lys Val Val 20 25 ' 30

Gly His Asp Ser Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala 35 40 45

Glu Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys
50 55 60

Tyr Glu Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro 65 70 75 80

Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu Phe Arg Asp Ser 85 90 95

Glu Leu Ile Glu Glu Asn Tyr Leu Gly Val Arg Asn Ala Ala Gln Arg 100 105 110

Lys Leu Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp 115 120 125

Glu Phe Thr Pro Leu Gly Arg Ile Leu Tyr Lys Ala Pro Ser Asp Gly 130 135 140

Lys Trp Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Ile Val Arg Asp 145 150 155 160

Val Lys Tyr Asp Pro Asn Pro Asp Glu Val Ala Asp Ala Lys Tyr Val 165 170 175

Asn Arg Glu Glu Leu Arg Glu Ile Leu Arg Lys Ala Asp Ala Gly Glu 180 185 190

Glu Gly Leu Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe 195 200 205

Leu Phe Lys Trp Trp Asp His Val Glu Gln Gly Thr Ile Lys Glu Val 210 215 220

Ala Asp Met Lys Thr Ile His Lys Leu Thr 225 230

<210> 41 <211> 238

÷ 1.



<212> PRT <213> Oryza Sativa

Gln Lys Arg Leu Met Phe Asp Asp Glu Cys Ile Leu Val Asp Glu Gln
20 25 30

Asp Asn Val Val Gly His Glu Ser Lys Tyr Asn Cys His Leu Met Glu 35 40 45

Lys Ile Glu Ser Glu Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu 50 60

Phe Asn Ser Lys Tyr Glu Leu Leu Gln Gln Arg Ser Ala Thr Lys
65 70 75 80

Val Thr Phe Pro Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu 85 90 95

Tyr Arg Glu Ser Glu Leu Ile Gln Glu Asn Tyr Leu Gly Val Arg Asn 100 105 110

Ala Ala Gln Arg Lys Leu Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp 115 120 125

Val Pro Val Asp Gln Phe Thr Pro Leu Gly Arg Met Leu Tyr Lys Ala 130 135 140

Pro Ser Asp Gly Lys Trp Gly Glu His Glu Leu Asp Tyr Leu Leu Phe 145 150 155 160

Ile Val Arg Asp Val Lys Val Val Pro Asn Pro Asp Glu Val Ala Asp 165 170 175

Val Lys Tyr Val Ser Arg Glu Gln Leu Lys Glu Leu Ile Arg Lys Ala 180 . 185 . 190

Asp Ala Gly Glu Glu Leu Lys Leu Ser Pro Trp Phe Arg Leu Val 195 200 205

Val Asp Asn Phe Leu Met Gly Trp Trp Asp His Val Glu Lys Gly Thr 210 215 220

Leu Asn Glu Ala Val Asp Met Glu Thr Ile His Lys Leu Lys 225 230 235

<210> 42

<211> 233

<212> PRT

<213> Arabidopsis thaliana

<400> 42

Met Thr Asp Ser Asn Asp Ala Gly Met Asp Ala Val Gln Arg Arg Leu
1 5 10 15

Met Phe Glu Asp Glu Cys Ile Leu Val Asp Glu Asn Asn Arg Val Val 20 25 30



- Gly His Asp Thr Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala 35 40 45
- Glu Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys 50 55 60
- Tyr Glu Leu Leu Gln Gln Arg Ser Lys Thr Lys Val Thr Phe Pro 65 70 75 80
- Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser 85 90 95
- Glu Leu Ile Glu Glu Asn Val Leu Gly Val Arg Asn Ala Ala Gln Arg 100 105 110
- Lys Leu Phe Asp Glu Leu Gly Ile Val Ala Glu Asp Val Pro Val Asp 115
- Glu Phe Thr Pro Leu Gly Arg Met Leu Tyr Lys Ala Pro Ser Asp Gly 130 140
- Lys Trp Gly Glu His Glu Val Asp Tyr Leu Leu Phe Ile Val Asp Asp 145 150 155 160
- Val Lys Leu Gln Pro Asn Pro Asp Glu Val Ala Glu Ile Lys Tyr Val 165 170 175
- Ser Arg Glu Glu Leu Lys Glu Leu Val Lys Lys Ala Asp Ala Gly Asp 180 185 190
- Glu Ala Val Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe 195 200 205
- Leu Met Lys Trp Trp Asp His Val Glu Lys Gly Thr Ile Thr Glu Ala 210 215 220
- Ala Asp Met Lys Thr Ile His Lys Leu 225 230
- <210> 43
- <211> 293
- <212> PRT
- <213> Haematococcus pluvialis
- <400> 43
- Met Leu Arg Ser Leu Leu Arg Gly Leu Thr His Ile Pro Arg Val Asn 1 5 10 15
- Ser Ala Gln Gln Pro Ser Cys Ala His Ala Arg Leu Gln Phe Lys Leu 20 25 30
- Arg Ser Met Gln Leu Leu Ser Glu Asp Arg Thr Asp His Met Arg Gly 35 40 45
- Ala Ser Thr Trp Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys 50 55 60
- Asp Glu Cys Ile Leu Val Asp Val Glu Asp Asn Ile Thr Gly His Ala 65 70 75 80
- Ser Lys Leu Glu Cys His Lys Phe Leu Pro His Gln Pro Ala Gly Leu

Leu His Arg Ala Phe Ser Val Phe Leu Phe Asp Asp Gln Gly Arg Leu 100 105 110

Leu Leu Gln Gln Arg Ala Arg Ser Lys Ile Thr Phe Pro Ser Val Trp 115 120 125

Thr Asn Thr Cys Cys Ser His Pro Leu His Gly Gln Thr Pro Asp Glu 130 135 140

Val Asp Gln Leu Ser Gln Val Ala Asp Gly Thr Val Pro Gly Ala Lys 145 150 155 160

Ala Ala Ala Ile Arg Lys Leu Glu His Glu Leu Gly Ile Pro Ala His 165 170 175

Gln Leu Pro Ala Ser Ala Phe Arg Phe Leu Thr Arg Leu His Tyr Cys 180 185 190

Ala Ala Asp Val Gln Pro Ala Ala Thr Gln Ser Ala Leu Trp Gly Glu 195 200 205

His Glu Met Asp Tyr Ile Leu Phe Ile Arg Ala Asn Val Thr Leu Ala 210 215 220

Pro Asn Pro Asp Glu Val Asp Glu Val Arg Tyr Val Thr Gln Glu Glu 225 230 235 240

Leu Arg Gln Met Met Gln Pro Asp Asn Gly Leu Gln Trp Ser Pro Trp 245 250 255

Phe Arg Ile Ile Ala Ala Arg Phe Leu Glu Arg Trp Trp Ala Asp Leu 260 265 270

Asp Ala Ala Leu Asn Thr Asp Lys His Glu Asp Trp Gly Thr Val His

His Ile Asn Glu Ala 290

<210> 44

<211> 304

<212> PRT

<213> Haematococcus pluvialis

<400> 44

Met Leu Arg Ser Leu Leu Arg Gly Leu Thr His Ile Pro Arg Val Asn 1 5 10 15

Ser Ala Gln Gln Pro Ser Cys Ala His Ala Arg Leu Gln Phe Lys Leu 20 25 30

Arg Ser Met Gln Met Thr Leu Met Gln Pro Ser Ile Ser Ala Asn Leu 35 40 45

Ser Arg Ala Glu Asp Arg Thr Asp His Met Arg Gly Ala Ser Thr Trp 50 55 60

Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys Asp Glu Cys Ile 65 70 75 80

WO 99/63055



- Leu Val Asp Val Glu Asp Asn Ile Thr Gly His Ala Ser Lys Leu Glu 85 90 95
- Cys His Lys Phe Leu Pro His Pro Ala Gly Leu Leu His Arg Ala Phe 100 105 110
- Ser Val Phe Leu Phe Asp Asp Gln Gly Arg Leu Leu Gln Gln Arg 115 120 125
- Ala Arg Ser Lys Ile Thr Phe Pro Ser Val Trp Thr Asn Thr Cys Cys 130 135
- Ser His Pro Leu His Gly Gln Thr Pro Asp Glu Val Asp Gln Leu Ser 145 150 155 160
- Gln Val Ala Asp Gly Thr Val Pro Gly Ala Lys Ala Ala Ala Ile Arg 165 170 175
- Lys Leu Glu His Glu Leu Gly Ile Pro Ala His Gln Leu Pro Ala Ser 180 185 190
- Ala Phe Arg Phe Leu Thr Arg Leu His Tyr Cys Ala Ala Asp Val Gln
 195 200 205
- Pro Ala Ala Thr Gln Ser Ala Leu Trp Gly Glu His Glu Met Asp Tyr 210 215 220
- Ile Leu Phe Ile Arg Ala Asn Val Thr Leu Ala Pro Asn Pro Asp Glu 225 230 235 240
- Val Asp Glu Val Arg Tyr Val Thr Gln Glu Glu Leu Arg Gln Met Met
 245 250 255
- Gln Pro Asp Asn Gly Leu Gln Trp Ser Pro Trp Phe Arg Ile Ile Ala 260 265 270
- Ala Arg Phe Leu Glu Arg Trp Trp Ala Asp Leu Asp Ala Ala Leu Asn 275 280 285
- Thr Asp Lys His Glu Asp Trp Gly Thr Val His His Ile Asn Glu Ala 290 295 300
- <210> 45
- <211> 307
- <212> PRT
- <213> Chlamydomonas reinhardtii
- <400> 45
- Met Arg Ser Ser Phe Ile Glu Pro Lys Pro Arg Ala Gln Pro Val Leu 1 15
- Ser Arg Gly Arg Ala Ser Met Arg Leu Ala Gln Ser Arg Ala Leu Val 20 25 30
- Ala Arg Val Ser Ser Ala Leu Trp Pro Gly Ala Gly Leu Ser Gln Ala 35 40 45
- Gln Ser Val Ala Val Arg Met Ala Ser Ser Ser Thr Trp Glu Gly Thr 50 55 60
- Gly Leu Ser Gln Asp Asp Phe Met Gln Arg Asp Glu Cys Leu Val Val

WO 99/63055

| 65 | | | | | 70 | | | | | 7,5 | | | | | 80 | | |
|--------------|------------------------------|------------|------------|---------------|------------|------------|---------------|------------|------------|------------|------------|------------|------------|------------|------------|----|---|
| Asp | Glu | Gln | Asp | Arg 85 | Leu | Leu | Gly | Thr | Ala 90 | Asn | Lys | Tyr | Asp | Cys 95 | His | | |
| Arg | Phe | Glu | Ala 100 | Ala | Lys | Gly | Gln | Pro 105 | Cys | Gly | Arg | Leu | His 110 | Arg | Ala | | |
| Phe | Ser | Val 115 | Phe | Leu | Phe | Ser | Pro 120 | Asp | Gly | Arg | Leu | Leu 125 | Leu | Gln | Gln | | |
| Arg | Ala 130 | Ala | Ser | Lys | Val | Thr 135 | Phe | Pro | Gly | Val | Trp 140 | Thr | Asn | Thr | Cys | | |
| Cys 145 | Ser | His | Pro | Leu | Ala 150 | Gly | Gln | Ala | Pro | Asp 155 | Glu | Val | Asp | Leu | Pro 160 | | |
| Ala | Ala | Val | Ala | Ser 165 | Gly | Gln | Val | Pró | Gly 170 | Ile | Lys | Ala | Ala | Ala 175 | Val | | |
| Arg | Lys | Leu | Gln 180 | His | Glu | Leu | Gly | Ile 185 | Pro | Pro | Glu | Gln | Val 190 | Pro | Ala | | |
| Ser | Ser | Phe 195 | Ser | Phe | Leu | Thr | Arg 200 | Leu | His | Tyr | Cys | Ala 205 | Ala | Asp | Thr | | |
| Ala | Thr 210 | His | Gly | Pro | Ala | Ala 215 | Glu | Trp | Gly | Glu | His 220 | | Val | Asp | Tyr | | |
| Val 225 | Leu | Phe | Val | Arg | Pro 230 | Gln | Gln | Pro | Val | Ser 235 | | Gln | Pro | Asn | Pro 240 | | |
| Asp | Glu | Val | Asp | Ala 245 | | Arg | Tyr | Val | Thr 250 | Leu | Pro | Glu | Leu | Gln 255 | Ser | | |
| Met | Met | Ala | Asp 260 | | Gly | Leu | Ser | Trp 265 | Ser | Pro | Trp | Phe | Arg 270 | Ile | Leu , | | |
| Ala | Thr | Gln 275 | | Ala | Phe | Leu | Pro 280 | | Trp | Trp | Gly | Asp 285 | | Lys | Arg | | |
| Arg | Trp 290 | | Pro | Gly | Gly | Ser 295 | | Leu | Ser | Asp | 300 | | Thr | Ile | His | | |
| Arg 305 | Val | Met | • | | | | | | | | | | | | | | |
| <21 <21 | 0> 4 1> 1 2> D 3> A | 848 NA | is pa | alaes | stina | ı | | | | | | | | | | | |
| < 4 0 GAG | 0> 4 GAGAA | 6 AAA | GAG | rgtt <i>i</i> | ATA | 'AAT' | rgtt <i>i</i> | AC TO | STCG | CATTO | C TTO | CAA | CACA | TAT | rcagact | 6 | C |
| CCA | TTTI | CTT | GTT: | TCT | CTT (| CAAA | ACAA | CA A | ACTA | ATGT | G AC | GGAG' | PATC | TAG | CTATGGA | 12 | (|
| ACI | ACTI | GGT | GTT | CGCA | ACC : | CAT | CTCT' | rc T | rgcc | CTGT | C TG | GACT' | ГТТG | GAA | CAAGAAA | 18 | (|
| CCI | TAGT | TAGT | TCA | AAAC' | TAG (| CTTA' | TAAC | AT A | CATC | GATA' | T GG | TTCT' | TCTT | GTA | GAGTAGA | 24 | (|
| | | | | | | | | | | | | | | | | | |

WO 99/63055

| | | | | | 1 - 2 2 | |
|------------|--------------|-------------|--------------|-------------|--------------------|------|
| TTTTCAAGTG | AGGGCTGATG | GTGGAAGCGG | GAGTAGAACT | TCTGTTGCTT | | 300 |
| TTTTGTGGAC | GAGGAGGATT | TTATCAAAGC | TGGTGGTTCT | GAGCTTTTGT | TTGTCCAAAT | 360 |
| GCAGCAAACA | AAGTCTATGG | AGAAACAGGC | CAAGCTCGCC | GATAAGTTGC | CACCAATACC | 420 |
| TTTCGGAGAA | TCTGTGATGG | ACTTGGTTGT | AATAGGTTGT | GGACCTGCTG | GTCTTTCACT | 480 |
| GGCTGCAGAA | GCTGCTAAGC | TAGGCTTGAA | AGTTGGCCTT | ATTGGTCCTG | ATCTTCCTTT. | 540 |
| TACAAATAAT | TATGGTGTGT | GGGAAGACGA | GTTCAAAGAT | CTTGGACTTG | AACGTTGTAT | 600. |
| CGAGCATGCT | TGGAAGGACA | CCATCGTATA | TCTTGACAAT | GATGCTCCTG | TCCTTATTGG | 660 |
| TCGTGCATAT | GGACGAGTTA | GCCGGCATTT | GCTGCATGAA | GAGTTGCTGA | AAAGGTGTGT | 720 |
| CGAGTCAGGT | GTATCATATC | TGAATTCTAA | AGTGGAAAGG | ATCACTGAAG | CTGGTGATGG | 780 |
| CCATAGTCTT | GTAGTTTGTG | AAAACGACAT | CTTTATCCCT | TGCAGGCTTG | CTACTGTTGC | 840 |
| ATCTGGAGCA | GCTTCAGGGA | AACTTTTGGA | GTATGAAGTA | GGTGGCCCTC | GTGTTTGTGT | 900 |
| CCAAACTGCI | TATGGTGTGG | AGGTTGAGGT | GGAGAACAAT | CCATACGATC | CCAACTTAAT | 960 |
| GGTATTTATO | GACTACAGAG | ACTATATGCA | ACAGAAATTA | CAGTGCTCGG | AAGAAGAATA | 1020 |
| TCCAACATTI | CTCTATGTCA | TGCCCATGTC | GCCAACAAGA | СТТТТТТТТ | AGGAAACCTG | 1080 |
| TTTGGCCTCA | AAAGATGCCA | TGCCTTTCGA | TCTACTGAAG | AGAAAACTAA | TGTCACGATT | 1140 |
| GAAGACTCT | G GGTATCCAAC | TTACAAAAA | TTATGAAGAG | GAATGGTCTT | ATATTCCTGT | 1200 |
| TGGGGGTTC | r TTACCAAAC | A CAGAGCAAA | A GAACCTAGCA | TTTGGTGCT | CAGCAAGCAT | 1260 |
| GGTGCATCC | A GCAACAGGC | r ATTCGGTTG | r ACGATCACTA | TCAGAAGCT | CAAAATATGC | 1320 |
| TTCTGTAAT | r GCAAAGATT | r TGAAGCAAG | A TAACTCTGC | ATATGTGGTT | r CTGGACAAAG | 1380 |
| CAGTGCAGT | A AACATTTCA | A TGCAAGCAT | G GAGCAGTCT | TGGCCAAAG | G AGCGAAAACG | 1440 |
| TCAAAGAGC | A TTCTTTCTT | T TCGGGTTAG | A GCTTATTGT | G CAGCTAGAT | A TTGAAGCAAC | 1500 |
| CAGAACGTT | C TTTAGAACC | T TCTTCCGCT | T GCCAACTTG | ATGTGGTGG | G GTTTCCTTGG | 1560 |
| GTCTTCACT | A TCATCTTTC | G ATCTTGTAT | T GTTTTCCAT | G TACATGTÍT | G TTTTGGCCCC | 1620 |
| GAACAGCAT | G AGGATGTCA | C TŢGTGAGAC | A TTTGCTTTC | A GATCCTTCT | G GTGCAGTTAT | 1680 |
| GGTTAAAGC | T TACCTCGAA | A GGTAATCTG | T TTTATGAAA | C TATAGTGTC | T CATTAAATAA | 1740 |
| ATGAGGATO | C TTCGTATAT | G TATATGATC | A TCTCTATGT | а татсстата | T TCTAATCTCA | 1800 |
| TAAAGTAAT | C GAAAATTCA | TGATAGAAA | AAAAAAA A | AAAAAAA A | | 1848 |
| | | | | | | |

<210> 47

Met Glu Leu Gly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp

<211> 529

<212> PRT

^{. &}lt;213> Adonis palaestina

<400> 47

| | | | | - | | | | | | | | | | | 7.2 |
|-----------------|------------|------------|------------|--------------|------------|------------|------------|-------------|------------|------------|------------|------------|--------------|------------|------------|
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Thr | Phe | Gly | Thr 20 | Arg | Asn | Leu | Ser | Ser 25 | Ser | Lys | Leu | Ala | Туг 30 | Asn | lle |
| His | Arg | Tyr 35 | Gly | Ser | Ser | Cys | Arg 40 | Val | Asp | Phe | Gln | Val 45 | Arg | Ala | Asp |
| Gly | Gly 50 | Ser | Gly | Ser | Arg | Ser 55 | Ser | Val | Ala | Tyr | Lys 60 | Glu | Gly | Phe | Val |
| Asp 65 | Glu | Glu | Asp | Phe | Ile 70 | Lys | Ala | Gly | Gly | Ser 75 | Glu | Leu | Leu | Phe | Val 80 |
| Gln | Met | Gln | Gln | Thr 85 | Lys | Ser | Met | Glu | Lys 90 | Gln | Ala | Lys | Leu | Ala 95 | Asp |
| Lys | Leu | Pro | Pro 100 | Ile | Pro | Phe | Gly | Glu 105 | Ser | Val | Met | Asp | Leu 110 | Val | Val |
| Ile | Gly | Cys 115 | Gly | Pro | Ala | Gly | Leu 120 | Ser | Leu | Ala | Ala | Glu 125 | Ala | Ala | Lys |
| Leu | Gly 130 | | Lys | Val | Gly | Leu 135 | Ile | Gly | Pro | Asp | Leu 140 | Pro | Phe | Thr | Asn |
| Asn 145 | Tyr | Gly | Val | Trp | Glu 150 | | Glu | Phe | Lys | Asp 155 | Leu | Gly | Leu | Glu | Arg 160 |
| Cys | Ile | Glu | His | Ala 165 | | Lys | Asp | Thr | Ile 170 | Val | Tyr | Leu | Asp | Asn 175 | Asp |
| Ala | Pro | Val | Leu 180 | | Gly | Arg | Ala | Tyr 185 | Gly | Arg | Val | Ser | Arg 190 | His | Leu |
| Leu | His | Glu 195 | | Lev | Lev | Lys | 200 | Cys | : Val | Glu | Ser | Gly 205 | v Val | Ser | Tyr , |
| Leu | Asp 210 | | Lys | s Val | Glu | Arc 215 | | Thr | Glu | ı Ala | Gly 220 | / Asp | Gly | His | Ser |
| Leu 225 | | Val | L Cys | s Glu | Asr 230 | | ı Ile | e Phe | e Ile | 235 | Cys | s Arg | g Lev | a Ala | 240 |
| Val | . Ala | a Sea | r Gl | y Ala 24! | | a Ser | r Gly | y Lys | 250 | u Lei O | Gl د | тул | r Glu | val 255 | L Gly 5 |
| Gl ₃ | y Pro | o Ar | g Va 26 | | s Vai | l Glı | n Th | r Ala 26 | а Ту: 5 | r Gl | y Vai | l Gl | u Val 270 | l Glu | ນ Vai |

Glu Asn Asn Pro Tyr Asp Pro Asn Leu Met Val Phe Met Asp Tyr Arg

Asp Tyr Met Gln Gln Lys Leu Gln Cys Ser Glu Glu Glu Tyr Pro Thr

Phe Leu Tyr Val Met Pro Met Ser Pro Thr Arg Leu Phe Phe Glu Glu

Thr Cys Leu Ala Ser Lys Asp Ala Met Pro Phe Asp Leu Leu Lys Arg

280

295

47

330

285

335

305



Lys Leu Met Ser Arg Leu Lys Thr Leu Gly Ile Gln Val Thr Lys Val 340 345 350

Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn 355 360 365

Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val His 370 375 380

Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys 385 390 395 400

Tyr Ala Ser Val Ile Ala Lys Ile Leu Lys Gln Asp Asn Ser Ala Tyr 405 410 415

Val Val Ser Gly Gln Ser Ser Ala Val Asn Ile Ser Met Gln Ala Trp 420 425 430

Ser Ser Leu Trp Pro Lys Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu 435 440 445

Phe Gly Leu Glu Leu Ile Val Gln Leu Asp Ile Glu Ala Thr Arg Thr 450 455 460

Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp Trp Gly Phe 465 470 475 480

Leu Gly Ser Ser Leu Ser Ser Phe Asp Leu Val Leu Phe Ser Met Tyr 485 490 495

Met Phe Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His 500 505 510

Leu Leu Ser Asp Pro Ser Gly Ala Val Met Val Arg Ala Tyr Leu Glu 515 520 525

Ara

<210> 48

<211> 378

<212> PRT

<213> Potato

<400> 48

Asp Glu Phe Lys Asp Leu Gly Leu Gln Ala Cys Ile Glu His Val Trp
1 5 10 15

Arg Asp Thr Ile Val Tyr Leu Asp Asp Asp Pro Ile Leu Ile Gly
20 25 30

Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu His Glu Glu Leu Leu 35 40 45

Lys Arg Cys Val Glu Ala Gly Val Leu Tyr Leu Asn Ser Lys Val Asp 50 55 60

Arg Ile Val Glu Ala Thr Asn Gly His Ser Leu Val Glu Cys Glu Gly 65 70 75 80

Asp Val Val Ile Pro Cys Arg Phe Val Thr Val Ala Ser Gly Ala Ala 85 90 95

Ser Gly Lys Phe Leu Gln Tyr Glu Leu Gly Gly Pro Arg Val Ser Val 100 105 110

Gln Thr Ala Tyr Gly Val Glu Val Glu Val Asp Asn Asn Pro Phe Asp 115 120 125

Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp Tyr Val Arg His Asp 130 135 140

Ala Gln Ser Leu Glu Ala Lys Tyr Pro Thr Phe Leu Tyr Ala Met Pro 145 150 155 160

Met Ser Pro Thr Arg Val Phe Phe Glu Glu Thr Cys Leu Ala Ser Lys 165 170 175

Asp Ala Met Pro Phe Asp Leu Leu Lys Lys Lys Leu Met Leu Arg Leu 180 185 190

Asn Thr Leu Gly Val Arg Ile Lys Glu Ile Tyr Glu Glu Glu Trp Ser 195 200 205

Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Thr Leu 210 215 220

Ala Phe Gly Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser 225 230 235 240

Val Val Arg Ser Leu Ser Glu Ala Pro Lys Cys Ala Phe Val Leu Ala 245 250 255

Asn Ile Leu Arg Gln Asn His Ser Lys Asn Met Leu Thr Ser Ser Ser 260 265 270

Thr Pro Ser Ile Ser Thr Gln Ala Trp Asn Thr Leu Trp Pro Gln Glu 275 280 285

Arg Lys Arg Gln Arg Ser Phe Phe Leu Phe Gly Leu Ala Leu Ile Leu 290 295 300

Gln Leu Asp Ile Glu Gly Ile Arg Ser Phe Phe Arg Ala Phe Phe Arg 305 310 315 320

Val Pro Lys Trp Met Trp Gln Gly Phe Leu Gly Ser Ser Leu Ser Xaa 325 330 335

Ala Asp Leu Met Leu Phe Ala Phe Tyr Met Phe Ile Ile Ala Pro Asn 340 345 350

Asp Met Arg Arg Gly Leu Ile Arg His Leu Leu Ser Asp Pro Thr Gly 355 360 365

Ala Thr Leu Ile Arg Thr Tyr Leu Thr Phe 370 . 375

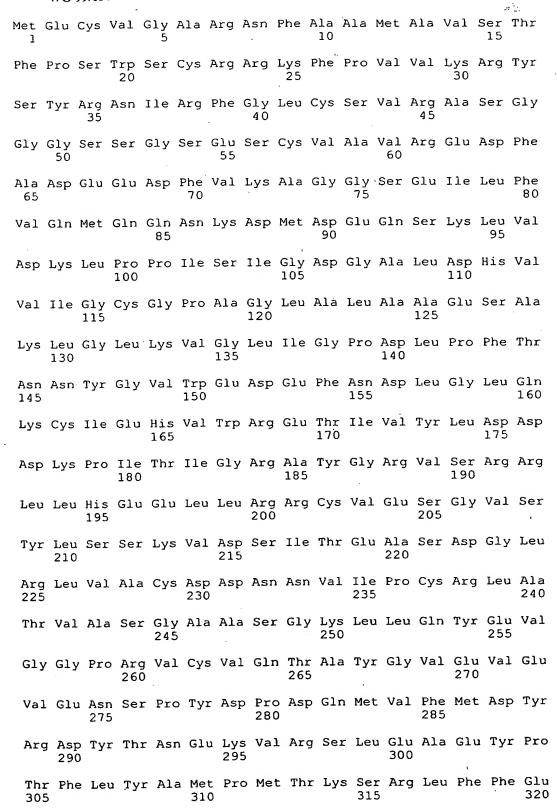
<210> 49

<211> 524

<212> PRT

<213> Arabidopsis thaliana

<400> 49



Glu Thr Cys Leu Ala Ser Lys Asp Val Met Pro Phe Asp Leu Leu Lys

325 330 335

Thr Lys Leu Met Leu Arg Leu Asp Thr Leu Gly Ile Arg Ile Leu Lys 340 345 350

Thr Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro 355 360 365

Asn Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ser Met Val $370 \hspace{1.5cm} 375 \hspace{1.5cm} 380$

His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro-385 390 395 400

Lys Tyr Ala Ser Val Ile Ala Glu Ile Leu Arg Glu Glu Thr Thr Lys
405
410
415

Gln Ile Asn Ser Asn Ile Ser Arg Gln Ala Trp Asp Thr Leu Trp Pro
420 425 430

Pro Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu Phe Gly Leu Ala Leu 435 440 445

Ile Val Gln Phe Asp Thr Glu Gly Ile Arg Ser Phe Phe Arg Thr Phe 450 455 460

Phe Arg Leu Pro Lys Trp Met Trp Gln Gly Phe Leu Gly Ser Thr Leu 465 470 475 480

Thr Ser Gly Asp Leu Val Leu Phe Ala Leu Tyr Met Phe Val Ile Ser 485 490 495

Pro Asn Asn Leu Arg Lys Gly Leu Ile Asn His Leu Ile Ser Asp Pro 500 505 510

Thr Gly Ala Thr Met Ile Lys Thr Tyr Leu Lys Val

<210> 50

<211> 529

<212> PRT

<213> Adonis palaestina

<400> 50

Met Glu Leu Cly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp 1 5 . 10 15

Thr Phe Gly Thr Arg Asn Leu Ser Ser Ser Lys Leu Ala Tyr Asn Ile 20 25 30

His Arg Tyr Gly Ser Ser Cys Arg Val Asp Phe Gln Val Arg Ala Asp 35 40 45

Gly Gly Ser Gly Ser Arg Ser Ser Val Ala Tyr Lys Glu Gly Phe Val
50 55 60

Asp Glu Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu Leu Leu Phe Val
65 70 75 80

Gln Met Gln Gln Thr Lys Ser Met Glu Lys Gln Ala Lys Leu Ala Asp

| Lys | Leu | Pro | Pro 100 | Ile | Pro | Phe | Gly | Glu 105 | Ser | Val | Met | Asp | Leu 110 | Val | Val |
|------------|------------|------------|------------|------------|------------|------------|------------|-------------|------------|------------|------------|------------|------------|------------|------------|
| Ile | Gly | Cys 115 | Gly | Pro | Ala | Gly | Leu 120 | Ser | Leu | Ala | Ala | Glu 125 | Ala | Ala | Lys |
| Leu | Gly 130 | Leu | Lys | Val | Gly | Leu 135 | Ile | Gly | Pro | Asp | Leu 140 | Pro | Phe | Thr | Asn |
| Asn 145 | Tyr | Gly | Val | Trp | Glu 150 | Asp | Glu | Phe | Lys | Asp 155 | | Gly | Leu | Glu | Arg 160 |
| Cys | Ile | Glu | His | Ala 165 | Trp | Lys | Asp | Thr | Ile 170 | Val | Tyr | Leu | Asp | Asn 175 | Asp |
| Ala | Pro | Val | Leu 180 | Ile | Gly | Arg | Ala | Tyr, 185 | Gly | Arg | Val | Ser | Arg 190 | His | Leu |
| Leu | His | Glu 195 | Glu | Leu | Leu | Lys | Arg 200 | Cys | Val | Glu | Ser | Gly 205 | Val | Ser | Tyr |
| Leu | Asp 210 | Ser | Lys | Val | Glu | Arg 215 | Ile | Thr | Glu | Ala | Gly 220 | Asp | Gly | His | Ser |
| Leu 225 | Val | Val | Cys | Glu | Asn 230 | Glu | Ile | Phe | Ile | Pro 235 | Cys | Arg | Leu | Ala | Thr 240 |
| Val | Ala | Ser | Gly | Ala 245 | Ala | Ser | Gly | Lys | Leu 250 | Leu | Glu | Tyr | Glu | Val 255 | Gly |
| Gly | Pro | Arg | Val 260 | Cys | Val | Gln | Thr | Ala 265 | Tyr | Gly | Val | Glu | Val 270 | Glu | Val |
| Glu | Asn | Asn 275 | Pro | Tyr | Asp | Pro | Asn 280 | Leu | Met | Val | Phe | Met 285 | Asp | Tyr | Arg |
| Asp | Tyr 290 | Met | Gln | Gln | Lys | Leu 295 | Gln | Cys | Ser | Glu | Glu 300 | Glu | Tyr | Pro | Thr |
| Phe 305 | Leu | Tyr | Val | Met | Pro 310 | Met | Ser | Pro | Thr | Arg 315 | Leu | Phe | Phe | Glu | Glu 320 |
| Thr | Cys | Leu | Ala | Ser 325 | Lys | Asp | Ala | Met | Pro 330 | Phe | Asp | Leu | Leu | Lys 335 | Arg |
| Lys | Leu | Met | Ser 340 | Arg | Leu | Lys | Thr | Leu 345 | Gly | Ile | Gln | Val | Thr 350 | Lys | Val |
| Tyr | Glu | Glu 355 | Glu | Trp | Ser | Tyr | Ile 360 | Pro | Val | Gly | Gly | Ser 365 | Leu | Pro | Asn |
| Thr | Glu 370 | Gln | Lys | Asn | Leu | Ala 375 | Phe | Gly | Ala | Ala | Ala 380 | Ser | Met | Val | His |
| Pro 385 | Ala | Thr | Gly | Tyr | Ser 390 | Val | Val | Arg | Ser | Leu 395 | Ser | Glu | Ala | Pro | Lys 400 |
| Tyr | Ala | Ser | Val | Ile 405 | Ala | Lys | Ile | Leu | Lys 410 | Gln | Asp | Asn | Ser | Ala 415 | Tyr |
| | | | | | | | | | | | | | | | |

Val Val Ser Gly Gln Ser Ser Ala Val Asn Ile Ser Met Gln Ala Trp 420 425 430

Ser Ser Leu Trp Pro Lys Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu 435 440 445

Phe Gly Leu Glu Leu Ile Val Gln Leu Asp Ile Glu Ala Thr Arg Thr 450 455 460

Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp Trp Gly Phe 465 470 480

Leu Gly Ser Ser Leu Ser Ser Phe Asp Leu Val Leu Phe Ser Met Tyr 485 490 495

Met Phe Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His 500 505 510

Leu Leu Ser Asp Pro Ser Gly Ala Val Met Val Arg Ala Tyr Leu Glu 515 520 525

Arg

<210> 51

<211> 529

<212> PRT

<213> Adonis palaestina

<400> 51

RNSDOCID: WO GORANESA 1 IA-

Met Glu Leu Gly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp

1 10 15

Thr Phe Gly Thr Arg Asn Leu Ser Ser Ser Lys Leu Ala Tyr Asn Ile 20 25 30

His Arg Tyr Gly Ser Ser Cys Arg Val Asp Phe Gln Val Arg Ala Asp 35 40 45

Gly Gly Ser Gly Ser Arg Thr Ser Val Ala Tyr Lys Glu Gly Phe Val 50 55 60

Asp Glu Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu Leu Leu Phe Val 65 70 75 80

Gln Met Gln Gln Thr Lys Ser Met Glu Lys Gln Ala Lys Leu Ala Asp 85 90 95

Lys Leu Pro Pro Ile Pro Phe Gly Glu Ser Val Met Asp Leu Val Val 100 105 110

Ile Gly Cys Gly Pro Ala Gly Leu Ser Leu Ala Ala Glu Ala Ala Lys 115 120 125

Leu Gly Leu Lys Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn 130 135 140

Asn Tyr Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Glu Arg 145 150 155 160



| 3055 | |
|------|-------|
| | _ |

Cys Ile Glu His Ala Trp Lys Asp Thr Ile Val Tyr Leu Asp Asn Asp 170 Ala Pro Val Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu His Glu Glu Leu Leu Lys Arg Cys Val Glu Ser Gly Val Ser Tyr Leu Asn Ser Lys Val Glu Arg Ile Thr Glu Ala Gly Asp Gly His Ser Leu Val Val Cys Glu Asn Asp Ile Phe Ile Pro Cys Arg Leu Ala Thr Val Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Glu Tyr Glu Val Gly Gly Pro Arg Val Cys Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val 265 Glu Asn Asn Pro Tyr Asp Pro Asn Leu Met Val Phe Met Asp Tyr Arg Asp Tyr Met Gln Gln Lys Leu Gln Cys Ser Glu Glu Glu Tyr Pro Thr 295 Phe Leu Tyr Val Met Pro Met Ser Pro Thr Arg Leu Phe Phe Glu Glu 310 Thr Cys Leu Ala Ser Lys Asp Ala Met Pro Phe Asp Leu Leu Lys Arg Lys Leu Met Ser Arg Leu Lys Thr Leu Gly Ile Gln Val Thr Lys Ile 345 Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn 360 Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys Tyr Ala Ser Val Ile Ala Lys Ile Leu Lys Gln Asp Asn Ser Ala Tyr Val Val Ser Gly Gln Ser Ser Ala Val Asn Ile Ser Met Gln Ala Trp Ser Ser Leu Trp Pro Lys Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu Phe Gly Leu Glu Leu Ile Val Gln Leu Asp Ile Glu Ala Thr Arg Thr Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp Trp Gly Phe 475 Leu Gly Ser Ser Leu Ser Ser Phe Asp Leu Val Leu Phe Ser Met Tyr

485

490

495

Met Phe Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His 500 505 510

Leu Leu Ser Asp Pro Ser Gly Ala Val Met Val Lys Ala Tyr Leu Glu 515 520 525

Arg

<210> 52

<211> 533

<212> PRT

<213> Lettuce

<400> 52

Met Glu Cys Phe Gly Ala Arg Asn Met Thr Ala Thr Met Ala Val Phe $1 \hspace{1cm} 5 \hspace{1cm} 10 \hspace{1cm} 15$

Thr Cys Pro Arg Phe Thr Asp Cys Asn Ile Arg His Lys Phe Ser Leu 20 25 30

Leu Lys Gln Arg Arg Phe Thr Asn Leu Ser Ala Ser Ser Ser Leu Arg 35 40 45

Gln Ile Lys Cys Ser Ala Lys Ser Asp Arg Cys Val Val Asp Lys Gln 50 55 60

Gly Ile Ser Val Ala Asp Glu Glu Asp Tyr Val Lys Ala Gly Gly Ser 65 70 75 80

Glu Leu Phe Phe Val Gln Met Gln Arg Thr Lys Ser Met Glu Ser Gln 85 90 95

Ser Lys Leu Ser Glu Lys Leu Ala Gln Ile Pro Ile Gly Asn Cys Ile 100 105 110

Leu Asp Leu Val Val Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala 115 120 125

Ala Glu Ser Ala Lys Leu Gly Leu Asn Val Gly Leu Ile Gly Pro Asp 130 135 140

Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Gln Asp Glu Phe Ile Gly 145 150 155 160

Leu Gly Leu Glu Gly Cys Ile Glu His Ser Trp Lys Asp Thr Leu Val 165 170 175

Tyr Leu Asp Asp Ala Asp Pro Ile Arg Ile Gly Arg Ala Tyr Gly Arg 180 185 190

Val His Arg Asp Leu Leu His Glu Glu Leu Leu Arg Arg Cys Val Glu 195 200 205

Ser Gly Val Ser Tyr Leu Ser Ser Lys Val Glu Arg Ile Thr Glu Ala 210 215 220

Pro Asn Gly Tyr Ser Leu Ile Glu Cys Glu Gly Asn Ile Thr Ile Pro

| | wo | 99/63 | 055 | | , | | | | | | | | | | -] |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 225 | | | | | 230 | | | | | 235 | | | | | 240 |
| Cys | Arg | Leu | Ala | Thr 245 | Val | Ala | Ser | Gly | Ala 250 | Ala | Ser | Gly | Lys | Phe 255 | Leu |
| Glu | Tyr | Glu | Leu 260 | Gly | Gly | Pro | Arg | Val 265 | Cys | Val | Gln | Thr | Ala 270 | Tyr | Gly |
| Ile | Glu | Val 275 | Glu | Val | Glu | Asn | Asn 280 | Pro | Tyr | Asp | Pro | Asp 285 | Leu | Met | Val |
| Phe | Met 290 | Asp | Tyr | Arg | Asp | Phe 295 | Ser | Lys | His | Lys | Pro 300 | Glu | Ser | Leu | Glu. |
| Ala 305 | Lys | Tyr | Pro | Thr | Phe 310 | Leu | Tyr | Val | Met | Ala 315 | Met | Ser | Pro | Thr | Lys 320 |
| Ile | Phe | Phe | Glu | Glu 325 | Thr | Cys | Leu | Ala | Ser 330 | Arg | Glu | Ala | Met | Pro 335 | Phe |
| Asn | Leu | Leu | Lys 340 | Ser | Lys | Leu | Met | Ser 345 | Arg | Leu | Lys | Ala | Met 350 | Gly | Ile |
| Arg | Ile | Thr 355 | Arg | Thr | Tyr | Glu | Glu 360 | Glu | Trp | Ser | Tyr | 11e 365 | Pro | Val | Gly |
| Gly | Ser 370 | Leu | Pro | Asn | Thr | Glu 375 | Gln | Lys | Asn | Leu | Ala 380 | Phe | Ġly | Ala | Ala |
| Ala 385 | Ser | Met | Val | His | Pro 390 | Ala | Thr | Gly | Tyr | Ser 395 | Val | Val | Arg | Ser | Leu 400 |
| Ser | Glu | Ala | Pro | Asn 405 | Tyr | Ala | Ala | Val | Ile 410 | Ala | Lys | Ile | Leu | Arg 415 | Gln |
| Asp | Gln | Ser | Lys 420 | Glu | Met | Ile | Ser | Leu 425 | Gly | Lys | Tyr | Thr | Asn 430 | Ile | Ser |
| Lys | Gln | Ala 435 | Trp | Glu | Thr | Leu | Trp 440 | Pro | Leu | Glu | Arg | Lys 445 | Arg | Gln | Arg |
| Ala | Phe 450 | Phe | Leu | Phe | Gly | Leu 455 | Ser | His | Ile | Val | Leu 460 | Met | Asp | Leu | Glu |
| Gly | Thr | Arg | Thr | Phe | Phe | Arg | Thr | Phe | Phe | Arg | Leu | Pro | Lvs | Trp | Met |

Thr Arg Thr Phe Phe Arg Thr Phe Phe Arg Leu Pro Lys Trp Met 470 475 480

Trp Trp Gly Phe Leu Gly Ser Ser Leu Ser Ser Thr Asp Leu Ile Ile 490

Phe Ala Leu Tyr Met Phe Val Ile Ala Pro His Ser Leu Arg Met Glu 505

Leu Val Arg His Leu Leu Ser Asp Pro Thr Gly Ala Thr Met Val Lys 515 520

Ala Tyr Leu Thr Ile 530

<210> 53

<211> 526 <212> PRT

<213> Tomato

<400> 53

Met Glu Cys Val Gly Val Gln Asn Val Gly Ala Met Ala Val Leu Thr 1 5 10 15

Arg Pro Arg Leu Asn Arg Trp Ser Gly Gly Glu Leu Cys Gln Glu Lys
20 25 30

Ser Ile Phe Leu Ala Tyr Glu Gln Tyr Glu Ser Lys Cys Asn Ser Ser 35 40 45

Ser Gly Ser Asp Ser Cys Val Val Asp Lys Glu Asp Phe Ala Asp Glu 50 55 60

Glu Asp Tyr Ile Lys Ala Gly Gly Ser Gln Leu Val Phe Val Gln Met 65 70 75 80

Gln Gln Lys Lys Asp Met Asp Gln Gln Ser Lys Leu Ser Asp Glu Leu 85 90 95

Arg Gln Ile Ser Ala Gly Gln Thr Val Leu Asp Leu Val Val Ile Gly 100 105 110

Cys Gly Pro Ala Gly Leu Ala Leu Ala Ala Glu Ser Ala Lys Leu Gly 115 120 125

Leu Asn Val Gly Leu Val Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr 130 135 140

Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Gln Ala Cys Ile 145 150 155 160

Glu His Val Trp Arg Asp Thr Ile Val Tyr Leu Asp Asp Asp Glu Pro 165 170 175

Ile Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Phe Leu His 180 185 190

Glu Glu Leu Lys Arg Cys Val Glu Ala Gly Val Leu Tyr Leu Asn 195 200 205

Ser Lys Val Asp Arg Ile Val Glu Ala Thr Asn Gly Gln Ser Leu Val 210 215 220

Glu Cys Glu Gly Asp Val Val Ile Pro Cys Arg Phe Val Thr Val Ala 225 230 235 240

Ser Gly Ala Ala Ser Gly Lys Phe Leu Gln Tyr Glu Leu Gly Ser Pro 245 250 255

Asn Pro Phe Asp Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp Tyr 275 280 285

Leu Arg His Asp Ala Gln Ser Leu Glu Ala Lys Tyr Pro Thr Phe Leu 290 295 300



| Tyr 305 | Ala | Met | Pro | Met | Ser 310 | Pro | Thr | Arg | Val | Phe 315 | Phe | Glu | Glu | Thr | Cys 320 |
|------------|------------|------------|------------|------------|------------|------------|---------------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Leu | Ala | Ser | Lys | Asp 325 | Ala | Met | Pro | Phe | Asp 330 | Leu | Leu | Lys | Lys | Lys 335 | Leu |
| Met | Leu | Arg | Leu 340 | Asn | Thr | Leu | Gly | Val 345 | Arg | Ile | Lys | Glu | 11e 350 | Tyr | Glu |
| Glu | Glu | Trp 355 | Ser | Tyr | Ile | Pro | Val 3 <u>6</u> 0 | Gly | Gly | Ser | Leu | Pro 365 | Asn | Thr | Glu |
| Gln | Lys 370 | Thr | Leu | Ala | Phe | Gly 375 | Ala | Ala | Ala | Ser | Met 380 | Val | His | Pro | Ala |
| Thr 385 | Gly | Tyr | Ser | Val | Val 390 | Arg | Ser | Leu | Ser | Glu 395 | Ala | Pro | Lys | Cys | Ala 400 |
| Ser | Val | Leu | Ala | Asn 405 | Ile | Leu | Arg | Gln | His 410 | Tyr | Ser | Lys | Asn | Met 415 | Leu |
| Thr | Ser | Ser | Ser 420 | Ile | Pro | Ser | Ile | Ser 425 | Thr | Gln | Ala | Trp | Asn 430 | Thr | Leu |
| Trp | Pro | Gln 435 | Glu | Arg | Lys | Arg | Gln 440 | Arg | Ser | Phe | Phe | Leu 445 | Phe | Gly | Leu |
| Ala | Leu 450 | lle | Leu | Gln | Leu | Asp 455 | Ile | Glu | Gly | Ile | Arg 460 | Ser | Phe | Phe | Arg |
| Ala 465 | Phe | Phe | Arg | Val | Pro 470 | Lys | Trp | Met | Trp | Gln 475 | Gly | Phe | Leu | | Ser 480 |
| Ser | Leu | Ser | Ser | Ala 485 | Asp | Leu | Met | Leu | Phe 490 | Ala | Phe | Tyr | Met | Phe 495 | Ile |
| Ile | Ala | Pro | Asn 500 | Asp | Met | Arg | Lys | Gly 505 | Leu | Ile | Arg | His | Leu 510 | Leu | Ser |
| Asp | Pro | Thr 515 | Gly | Ala | Thr | Leu | Ile 520 | Arg | Thr | Tyr | Leu | Thr 525 | Phe | | |

<210> 54

<211> 516

<212> PRT

<213> Tagetes erecta

<400> 54

BRICOCIO -MIO MOSACERE IA-

Met Ser Met Arg Ala Gly His Met Thr Ala Thr Met Ala Ala Phe Thr $1 \hspace{1cm} 5 \hspace{1cm} 10 \hspace{1cm} 15$

Cys Pro Arg Phe Met Thr Ser Ile Arg Tyr Thr Lys Gln Ile Lys Cys 20 25 30

Asn Ala Ala Lys Ser Gln Leu Val Val Lys Gln Glu Ile Glu Glu Glu 35 40 45

Glu Asp Tyr Val Lys Ala Gly Gly Ser Glu Leu Leu Phe Val Gln Met 50 55 60

BNSDOCID: <WO - 006305581 IAS

| Gln 65 | Gln | Asn | Lys | Ser | Met 70 | Asp | Ala | Gln | Ser | Ser 75 | Leu | Ser | Gln | Lys | Leu 80 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Pro | Arg | Val | Pro | Ile 85 | Gly | Gly | Gly | Gly | Asp 90 | Ser | Asn | Cys | Ile | Leu 95 | Asp |
| Leu | Val | Val | Ile 100 | Gly | Cys | Gly | Pro | Ala 105 | Gly | Leu | Ala | Leu | Ala 110 | Gly | Glu |
| Ser | Ala | Lys 115 | Leu | Gly | Leu | Asn | Val 120 | Ala | Leu | Ile | Gly | Pro 125 | Asp | Leu | Pro |
| Phe | Thr 130 | Așn | Asn | Tyr | Gly | Val 135 | Trp | Glu | Asp | Glu | Phe 140 | Ile | Gly | Leu | Gly |
| Leu 145 | Glu | Gly | Cys | Ile | Glu 150 | His | Val | Trp | Arg | Asp 155 | Thr | Val | Val | Tyr | Leu 160 |
| Asp | Asp | Asn | Asp | Pro 165 | Ile | Leu | Ile | Gly | Arg 170 | Ala | Tyr | Gly | Arg | Val 175 | Ser |
| Arg | Asp | Leu | Leu 180 | His | Glu | Glu | Leu | Leu 185 | Thr | Arg | Суѕ | Met | Glu 190 | Ser | Gly |
| Val | Ser | Tyr 195 | Leu | Ser | Ser | Lys | Val 200 | Glu | Arg | Ile | Thr | Glu 205 | Ala | Pro | Asn |
| Gly | Leu 210 | Ser | Leu | Ile | Glu | Cys 215 | Glu | Gly | Asn | Ile | Thr 220 | Ile | Pro | Cys | Arg |
| Leu 225 | Ala | Thr | Val | Ala | Ser 230 | Gly | Ala | Ala | Ser | Gly 235 | Lys | Leu | Leu | Gln | Tyr 240 |
| Glu | Leu | Gly | Gly | Pro 245 | Arg | Val | Cys | Val | Gln 250 | Thr | Ala | Tyr | Gly | Ile 255 | Glu |
| Val | Glu | Val | Glu 260 | Ser | Ile | Pro | Tyr | Asp 265 | Pro | Ser | Leu | Met | Val 270 | Phe | Met |
| Asp | Tyr | Arg 275 | Asp | Tyr | Thr | Lys | His 280 | Lys | Ser | Gln | Ser | Leu 285 | Glu | Ala | Gln |
| Tyr | Pro 290 | Thr | Phe | Leu | Tyr | Val 295 | Met | Pro | Met | Ser | Pro 300 | Thr | Lys | Val | Phe |
| Phe 305 | Glu | Glu | Thr | Cys | Leu 310 | Ala | Ser | Lys | Glu | Ala 315 | Met | Pro | Phe | Glu | Leu 320 |
| Leu | Lys | Thr | Lys | Leu 325 | Met | Ser | Arg | Leu | Lys 330 | Thr | Met | Gly | Ile | Arg 335 | Ile |
| Thr | Lys | Thr | Tyr 340 | Glu | Glu | Glu | Trp | Ser 345 | Tyr | Ile | Pro | Val | Gly 350 | Gly | Ser |
| Leu | Pro | Asn 355 | Thr | Glu | Gln | Lys | Asn 360 | Leu | Ala | Phe | Gly | Ala 365 | Ala | Ala | Ser |
| Met | Val 370 | His | Pro | Ala | Thr | Gly 375 | Tyr | Ser | Val | Val | Arg 380 | Ser | Leu | Ser | Glu |
| Ala | Pro | Asn | Tyr | Ala | Ala | Val | Ile | Ala | Lys | Ile | Leu | Gly | Lys | Gly | Asn |

PCT/US99/12121

390 395 Ser Lys Gln Met Leu Asp His Gly Arg Tyr, Thr Thr Asn Ile Ser Lys 405 410 Gln Ala Trp Glu Thr Leu Trp Pro Leu Glu Arg Lys Arg Gln Arg Ala 425 Phe Phe Leu Phe Gly Leu Ala Leu Ile Val Gln Met Asp Ile Glu Gly Thr Arg Thr Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp 455 Trp Gly Phe Leu Gly Ser Ser Leu Ser Ser Thr Asp Leu Ile Ile Phe 475 Ala Phe Tyr Met Phe Ile Ile Ala Pro His Ser Leu Arg Met Gly Leu Val Arg His Leu Leu Ser Asp Pro Thr Gly Gly Thr Met Leu Lys Ala 500 505 Tyr Leu Thr Ile 515 <210> 55 <211> 501 <212> PRT <213> Arabidopsis thaliana <400> 55 Met Asp Thr Leu Leu Lys Thr Pro Asn Lys Leu Asp Phe Phe Ile Pro Gln Phe His Gly Phe Glu Arg Leu Cys Ser Asn Asn Pro Tyr His Ser Arg Val Arg Leu Gly Val Lys Lys Arg Ala Ile Lys Ile Val Ser Ser Val Val Ser Gly Ser Ala Ala Leu Leu Asp Leu Val Pro Glu Thr Lys Lys Glu Asn Leu Asp Phe Glu Leu Pro Leu Tyr Asp Thr Ser Lys Ser Gln Val Val Asp Leu Ala Ile Val Gly Gly Pro Ala Gly Leu Ala Val Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile Asp 105

Pro Ser Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val Asp

Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Thr Thr Trp Ser

Gly Ala Val Val Tyr Val Asp Glu Gly Val Lys Lys Asp Leu Ser Arg

135

WO 99/63055

| 145 | 1 | | | | 150 | | | | | 155 | | | | | 160 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Pro | Туг | Gly | Arg | Val 165 | Asn | Arg | Lys | Gln | Leu 170 | Lys | Ser | Lys | Met | Leu 175 | Gln |
| Lys | Cys | Ile | Thr 180 | Asn | Gly | Val | Lys | Phe 185 | His | Gln | Ser | Lys | Val 190 | | Asn |
| Val | Val | His 195 | Glu | Glu | Ala | Asn | Ser 200 | Thr | Val | Val | Cys | Ser 205 | Asp | Gly | Val |
| Lys | Ile 210 | Gln | Ala | Ser | Val | Val 215 | Leu | Asp | Ala | Thr | Gly 220 | Phe | Ser | Arg | Cys |
| Leu 225 | Val | Gln | Tyr | Asp | Lys 230 | Pro | Tyr | Asn | Pro | Gly 235 | Tyr | Gln | Val | Ala | Tyr 240 |
| Gly | Ile | Val | Ala | Glu 245 | Val | Asp | Gly | His | Pro 250 | Phe | Asp | Val | Asp | Lys 255 | Met |
| Val | Phe | Met | Asp 260 | Trp | Arg | Asp | Lys | His 265 | Leu | Asp | Ser | Tyr | Pro 270 | Glu | Leu |
| Lys | Glu | Arg 275 | Asn | Ser | Lys | Ile | Pro 280 | Thr | Phe | Leu | Tyr | Ala 285 | Met | Pro | Phe |
| Ser | Ser 290 | Asn | Arg | Ile | Phe | Leu 295 | Glu | Glu | Thr | Ser | Leu 300 | Val | Ala | Arg | Pro |
| Gly 305 | Leu | Arg | Met | Glu | Asp 310 | Ile | Gln | Glu | Arg | Met 315 | Ala | Ala | Arg | Leu | Lys 320 |
| His | Leu | Gly | Ile | Asn 325 | Val | Lys | Arg | Ile | Glu 330 | Glu | Asp | Glu | Arg | Cys 335 | Val |
| Ile | Pro | Met | Gly 340 | Gly | Pro | Leu | Pro | Val 345 | Leu | Pro | Gln | Arg | Val 350 | Val | Gly |
| Ile | Gly | Gly 355 | Thr | Ala | Gly | Met | Val 360 | His | Pro | Ser | Thr | Gly 365 | Tyr | Met | Val |
| Ala | Arg 370 | Thr | Leu | Ala | Ala | Ala 375 | Pro | Ile | Val | Ala | Asn 380 | Ala | Ile | Val | Arg |
| Tyr 385 | Leu | Gly | Ser | Pro | Ser 390 | Ser | Asn | Ser | Leu | Arg 395 | Gly | Asp | Gln | Leu | Ser 400 |
| Ala | Glu | Val | Trp | Arg 405 | Asp | Leu | Trp | Pro | Ile 410 | Glu | Arg | Arg | Arg | Gln 415 | Arg |
| Glu | Phe | Phe | Cys 420 | Phe | Gly | Met | Asp | 11e 425 | Leu | Leu | Lys | Leu | Asp 430 | Leu | Asp |
| Ala | Thr | Arg 435 | Arg | Phe | Phe | Asp | Ala 440 | Phe | Phe | Asp | Leu | Gln 445 | Pro | His | Tyr |
| Trp | His 450 | Gly | Phe | Leu | Ser | Ser 455 | Arg | Leu | Phe | Leu | Pro 460 | Glu | Leu | Leu | Val |
| Phe 465 | Gly | Leu | Ser | Leu | Phe 470 | Ser | His | Ala | Ser | Asn 475 | Thr | Ser | Arg | Leu | Glu 480 |



Ile Met Thr Lys Gly Thr Val Pro Leu Ala Lys Met Ile Asn Asn Leu 485 490 495

Val Gln Asp Arg Asp 500

<210> 56

<211> 502

<212> PRT

<213> Adonis palaestina

<400> 56

Met Asp Thr Leu Leu Arg Thr His Asn Lys Leu Glu Leu Leu Pro Thr 1 5 10 15

Leu His Gly Phe Ala Glu Lys Gln His Leu Val Ser Thr Ser Lys Leu 20 25. 30

Gln Asn Gln Val Phe Arg Ile Ala Ser Arg Asn Ile His Pro Cys Arg 35 40 45

Asn Gly Thr Val Lys Ala Arg Gly Ser Ala Leu Leu Glu Leu Val Pro 50 55 60

Glu Thr Lys Lys Glu Asn Leu Glu Phe Asp Leu Pro Ala Tyr Asp Pro 65 70 . 75 80

Ser Arg Gly Ile Val Val Asp Leu Ala Val Val Gly Gly Pro Ala 85 90 95

Gly Leu Ala Ile Ala Gln Gln Val Ser Glu Ala Gly Leu Leu Val Cys 100 105 110

Ser Ile Asp Pro Ser Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val 115 120 125

Trp Val Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Thr 130 135 140

Thr Trp Ser Gly Ala Val Val Tyr Thr Asp Asp Asn Ser Lys Lys Tyr 145 150 155 160

Leu Asp Arg Pro Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys 165 170 175

Met Leu Gln Lys Cys Val Thr Asn Gly Val Lys Phe His Gln Ala Lys 180 185 190

Val Ile Lys Val Ile His Glu Glu Ser Lys Ser Leu Leu Ile Cys Asn 195 200 205

Asp Gly Ile Thr Ile Asn Ala Thr Val Val Leu Asp Ala Thr Gly Phe 210 215 220

Ser Arg Cys Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln 225 230 235 240

Val Ala Tyr Gly Ile Met Ala Glu Val Glu Glu His Pro Phe Asp Leu 245 250 255



- Asp Lys Met Leu Phe Met Asp Trp Arg Asp Ser His Leu Asn Glu Lys 260 265 270
- Leu Glu Leu Lys Asp. Lys Asn Arg Lys Ile Pro Thr Phe Leu Tyr Ala 275 280 285
- Met Pro Phe Ser Ser Thr Lys Ile Phe Leu Glu Glu Thr Ser Leu Val 290 295 300
- Ala Arg Pro Gly Leu Arg Phe Glu Asp Ile Gln Glu Arg Met Val Ala 305 310 315 320
- Arg Leu Lys His Leu Gly Ile Lys Val Lys Ser Ile Glu Glu Asp Glu 325
- Arg Cys Val Ile Pro Met Gly Gly Pro Leu Pro Val Leu Pro Gln Arg 340 345 350
- Val Val Gly Ile Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly 355 360 365
- Tyr Met Val Ala Arg Thr Leu Ala Ala Ala Pro Val Val Ala Lys Ser 370 380
- Ile Val Gln Tyr Leu Gly Ser Asp Arg Ser Leu Ser Gly Asn Glu Leu 385 390 395 400
- Ser Ala Glu Val Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Gln 405 410 415
- Arg Glu Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu 420 425 430
- Gln Gly Thr Arg Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro His 435 440 445
- Tyr Trp His Gly Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Leu 450 455 460
- Phe Phe Gly Leu Ser Leu Phe Ser His Ala Ser Asn Ala Ser Arg Ile 465 470 475 480
- Glu Ile Met Ala Lys Gly Thr Val Pro Leu Val Asn Met Met Asn Asn 485
- Leu Ile Gln Asp Thr Asp 500
- <210> 57
- <211> 498
- <212> PRT
- <213> Pepper
- <400> 57
- Met Asp Thr Leu Leu Arg Thr Pro Asn Asn Leu Glu Phe Leu His Gly
 1 10 15
- Phe Gly Val Lys Val Ser Ala Phe Ser Ser Val Lys Ser Gln Lys Phe 20 25 30



Gly Ala Lys Lys Phe Cys Glu Gly Leu Gly Ser Arg Ser Val Cys Val Lys Ala Ser Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr Lys Lys Glu Asn Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys Gly Val Val Val Asp Leu Ala Val Val Gly Gly Pro Ala Gly Leu Ala Val Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile Asp Pro 105 Asn Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp Ser Gly 135 Ala Ala Val Tyr Ile Asp Asp Lys Thr Thr Lys Asp Leu Asn Arg Pro Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys Met Met Gln Lys Cys Ile Leu Asn Gly Val Lys Phe His Gln Ala Lys Val Ile Lys Val 185 Ile His Glu Glu Ser Lys Ser Met Leu Ile Cys Asn Asp Gly Ile Thr 200 Ile Gln Ala Thr Val Val Leu Asp Ala Thr Gly Phe Ser Arg Ser Leu 215 Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln Val Ala Tyr Gly Ile Leu Ala Glu Val Glu Glu His Pro Phe Asp Val Asn Lys Met Val 250 Phe Met Asp Trp Arg Asp Ser His Leu Lys Asn Asn Val Glu Leu Lys 265 Glu Arg Asn Ser Arg Ile Pro Thr Phe Leu Tyr Ala Met Pro Phe Ser Ser Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg Pro Gly Leu Gly Met Asp Asp Ile Gln Glu Arg Met Val Ala Arg Leu Ser His 305 310 315 Leu Gly Ile Lys Val Lys Ser Ile Glu Glu Asp Glu His Cys Val Ile 330 Pro Met Gly Gly Pro Leu Pro Val Leu Pro Gln Arg Val Val Gly Ile Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met Val Ala

÷ []

355 360 365

Arg Thr Leu Ala Ala Ala Pro Val Val Ala Asn Ala Ile Ile Gln Tyr 370 375 380

Leu Ser Ser Glu Arg Ser His Ser Gly Asp Glu Leu Ser Ala Ala Val 385 390 395 400

Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Arg Gln Arg Glu Phe Phe
405 410 415

Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu Pro Ala Thr Arg. 420 425 430

Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp His Gly 435 440 445

Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Ile Val Phe Gly Leu 450 455 460

Ser Leu Phe Ser His Ala Ser Asn Thr Ser Arg Leu Glu Ile Met Thr 465 470 475 480

Lys Gly Thr Leu Pro Leu Val His Met Ile Asn Asn Leu Leu Gln Asp 485 490 495

Lys Glu

<210> 58

<211> 500

<212> PRT

<213> Tomato

<400> 58

Met Asp Thr Leu Leu Lys Thr Pro Asn Asn Leu Glu Phe Leu Asn Pro 1 5 10 15

His His Gly Phe Ala Val Lys Ala Ser Thr Phe Arg Ser Glu Lys His 20 25 30

His Asn Phe Gly Ser Arg Lys Phe Cys Glu Thr Leu Gly Arg Ser Val 35 40 45

Cys Val Lys Gly Ser Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr
50 .55 60

Lys Lys Glu Asn Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys
65 70 75 80

Gly Val Val Val Asp Leu Ala Val Val Gly Gly Gly Pro Ala Gly Leu 85 90 95

Ala Val Ala Gl
n Gl
n Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile 100
 $105\,$ $110\,$

Asp Pro Asn Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val 115 120 125

Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp



ان ما 135 140 Ser Gly Ala Ala Val Tyr Ile Asp Asp Asn Thr Ala Lys Asp Leu His 150 155 Arg Pro Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys Met Met Gln Lys Cys Ile Met Asn Gly Val Lys Phe His Gln Ala Lys Val Ile 185 Lys Val Ile His Glu Glu Ser Lys Ser Met Leu Ile Cys Asn Asp Gly 200 Ile Thr Ile Gln Ala Thr Val Val Leu Asp Ala Thr Gly Phe Ser Arg Ser Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln Val Ala 230 Tyr Gly Ile Leu Ala Glu Val Glu Glu His Pro Phe Asp Val Asn Lys 250 Met Val Phe Met Asp Trp Arg Asp Ser His Leu Lys Asn Asn Thr Asp 265 Leu Lys Glu Arg Asn Ser Arg Ile Pro Thr Phe Leu Tyr Ala Met Pro Phe Ser Ser Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg Pro Gly Leu Arg Ile Asp Asp Ile Gln Glu Arg Met Val Ala Arg Leu 310 315 Asn His Leu Gly Ile Lys Val Lys Ser Ile Glu Glu Asp Glu His Cys 330 Leu Ile Pro Met Gly Gly Pro Leu Pro Val Leu Pro Gln Arg Val Val 345 Gly Ile Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met 360 Val Ala Arg Thr Leu Ala Ala Ala Pro Val Val Ala Asn Ala Ile Ile Gln Tyr Leu Gly Ser Glu Arg Ser His Ser Gly Asn Glu Leu Ser Thr 395 Ala Val Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Gln Arg Glu Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu Pro Ala 425 Thr Arg Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp 440 His Gly Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Ile Val Phe

450

Li

Gly Leu Ser Leu Phe Ser His Ala Ser Asn Thr Ser Arg Phe Glu Ile 465 470 475 480

Met Thr Lys Gly Thr Val Pro Leu Val Asn Met Ile Asn Asn Leu Leu 485 490 495

Gln Asp Lys Glu 500

<210> 59

<211> 500

<212> PRT

<213> Tobacco

<400> 59

BNSDOCID: <WO

996305541 145

Met Asp Thr Leu Leu Lys Thr Pro Asn Lys Leu Glu Phe Leu His Pro 1 5 10 15

Val His Gly Phe Ser Val Lys Ala Ser Ser Phe Asn Ser Val Lys Pro 20 25 30

His Lys Phe Gly Ser Arg Lys Ile Cys Glu Asn Trp Gly Lys Gly Val 35 40 45

Cys Val Lys Ala Lys Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr 50 55 60

Lys Lys Glu Asn Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys 65 70 75 80

Gly Leu Val Val Asp Leu Ala Val Val Gly Gly Pro Ala Gly Leu 85 90 95

Ala Val Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Val Ser Ile 100 105 110

Asp Pro Ser Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val 115 120 125

Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp 130 135 140

Ser Gly Thr Val Val Tyr Ile Asp Asp Asn Thr Thr Lys Asp Leu Asp 145 150 155 160

Arg Pro Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys Met Met 165 170 175

Gln Lys Cys Ile Leu Asn Gly Val Lys Phe His His Ala Lys Val Ile 180 185 190

Lys Val Ile His Glu Glu Ala Lys Ser Met Leu Ile Cys Asn Asp Gly
195 200 205

Val Thr Ile Gln Ala Thr Val Val Leu Asp Ala Thr Gly Phe Ser Arg 210 215 220

Cys Leu Val Gln Tyr Asp Lys Pro Tyr Lys Pro Gly Tyr Gln Val Ala 225 230 235 240



Tyr Gly Ile Leu Ala Glu Val Glu Glu His Pro Phe Asp Thr Ser Lys 245

Met Val Leu Met Asp Trp Arg Asp Ser His Leu Gly Asn Asn Met Glu 260

Leu Lys Glu Arg Asn Arg Lys Val Pro Thr Phe Leu Tyr Ala Met Pro 275 280 285

Phe Ser Ser Asn Lys Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg 290 295 300

Pro Gly Leu Arg Met Asp Asp Ile Gln Glu Arg Met Val Ala Arg Leu 305 310 315 320

Asn His Leu Gly Ile Lys Val Lys Ser Ile Glu Glu Asp Glu His Cys 325 330 335

Val Ile Pro Met Gly Gly Ser Leu Pro Val Ile Pro Gln Arg Val Val 340 345 350

Gly Thr Gly Gly Thr Ala Gly Leu Val His Pro Ser Thr Gly Tyr Met 355 360 365

Val Ala Arg Thr Leu Ala Ala Ala Pro Val Val Ala Asn Ala Ile Ile 370 375 380

His Tyr Leu Gly Ser Glu Lys Asp Leu Leu Gly Asn Glu Leu Ser Ala 385 390 395 400

Ala Val Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Arg Gln Arg Glu

Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu Pro Ala 420 425 430

Thr Arg Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp 435 440 445

His Gly Phe Leu Ser Ser Arg Leu Tyr Leu Pro Glu Leu Ile Phe Phe 450 455 460

Gly Leu Ser Leu Phe Ser Arg Ala Ser Asn Thr Ser Arg Ile Glu Ile 465 470 475. 480

Met Thr Lys Gly Thr Leu Pro Leu Val Asn Met Ile Asn Asn Leu Leu 485 490 495

Gln Asp Thr Glu . 500

<210> 60

<211> 511

<212> PRT

<213> Tagetes erecta

<400> 60

Met Asp Thr Phe Leu Arg Thr Tyr Asn Ser Phe Glu Phe Val His Pro 1 5 10 15

BNISDOCID -WO

Ser Asn Lys Phe Ala Gly Asn Leu Asn Asn Leu Asn Gln Leu Asn Gln 20 25 30

Ser Lys Ser Gln Phe Gln Asp Phe Arg Phe Gly Pro Lys Lys Ser Gln 35 40 45

Phe Lys Leu Gly Gln Lys Tyr Cys Val Lys Ala Ser Ser Ser Ala Leu 50 . 55 60

Leu Glu Leu Val Pro Glu Ile Lys Lys Glu Asn Leu Asp Phe Asp Leu 65 70 75 80

Pro Met Tyr Asp Pro Ser Arg Asn Val Val Val Asp Leu Val Val Val Val Val 95

Gly Gly Gly Pro Ser Gly Leu Ala Val Ala Gln Gln Val Ser Glu Ala 100 105 110

Gly Leu Thr Val Cys Ser Ile Asp Pro Ser Pro Lys Leu Ile Trp Pro 115 120 125

Asn Asn Tyr Gly Val Trp Val Asp Glu Phe Glu Ala Met Asp Leu Leu 130 135 140

Asp Cys Leu Asp Thr Thr Trp Ser Ser Ala Val Val Tyr Ile Asp Glu 145 150 155 160

Lys Ser Thr Lys Ser Leu Asn Arg Pro Tyr Ala Arg Val Asn Arg Lys 165 170 175

Gln Leu Lys Thr Lys Met Leu Gln Lys Cys Ile Ala Asn Gly Val Lys 180 185 190

Phe His Gln Ala Lys Val Ile Lys Val Ile His Glu Glu Leu Lys Ser 195 200 205

Leu Leu Ile Cys Asn Asp Gly Val Thr Ile Gln Ala Thr Leu Val Leu 210 215 220

Asp Ala Thr Gly Phe Ser Arg Ser Leu Val Gln Tyr Asp Lys Pro Tyr 225 230 235 240

Asn Pro Gly Tyr Gln Val Ala Tyr Gly Ile Leu Ala Glu Val Glu Glu 245 250 255

His Pro Phe Asp Val Asp Lys Met Leu Phe Met Asp Trp Arg Asp Ser 260 265 270

His Leu Asp Gln Asn Leu Glu Ile Lys Ala Arg Asn Ser Arg Ile Pro 275 280 285

Thr Phe Leu Tyr Ala Met Pro Phe Ser Ser Thr Arg Ile Phe Leu Glu 290 295 300

Glu Thr Ser Leu Val Ala Arg Pro Gly Leu Lys Met Glu Asp Ile Gln 305 310 315 320

Glu Arg Met Ala Tyr Arg Leu Lys His Leu Gly Ile Lys Val Lys Ser 325 330 335

Ile Glu Glu Asp Glu Arg Cys Val Ile Pro Met Gly Gly Pro Leu Pro

340 345 350 Val Leu Pro Gln Arg Val Leu Gly Ile Gly Gly Thr Ala Gly Met Val 360 His Pro Ser Thr Gly Tyr Met Val Ala Arg Thr Leu Ala Ala Pro Ile Val Ala Lys Ser Ile Ile Arg Tyr Leu Asn Asn Glu Lys Ser Met Val Ala Asp Val Thr Gly Asp Asp Leu Ala Ala Gly Ile Trp Arg Glu Leu Trp Pro Ile Glu Arg Arg Gln Arg Glu Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu Glu Gly Thr Arg Arg Phe Phe 440 Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp His Gly Phe Leu Ser 455 Ser Arg Leu Phe Leu Pro Glu Leu Val Thr Phe Gly Leu Ser Leu Phe 470 Gly His Ala Ser Asn Thr Cys Arg Val Glu Ile Met Ala Lys Gly Thr 490 Leu Pro Leu Ala Thr Met Ile Gly Asn Leu Val Arg Asp Arg Glu 500 505

<210> 61

<211> 503

<212> PRT

<213> Daffodil

<400> 61

Met Asp Thr Leu Leu Arg Thr His Asn Arg Leu Glu Leu Leu Tyr Pro 1 5 10 15

Leu His Glu Leu Ala Lys Arg His Phe Leu Ser Pro Ser Pro Asn Pro 20 25 30

Gln Asn Pro Asn Phe Lys Phe Phe Ser Arg Lys Pro Tyr Gln Lys Lys 35 40 45

Cys Arg Asn Gly Tyr Ile Gly Val Ser Ser Asn Gln Leu Leu Asp Leu 50 55 60

Val Pro Glu Ile Lys Lys Glu His Leu Glu Phe Asp Leu Pro Leu Tyr 65 70 75 80

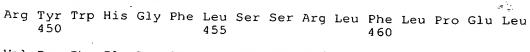
Asp Pro Ser Lys Ala Leu Thr Leu Asp Leu Ala Val Val Gly Gly 85 90 95

Pro Leu Ala Arg Ser Cys Ser Thr Ser Leu Gly Gly Gly Leu Ser Val

Val Ser Ile Asp Pro Asn Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly

BNSDOCID: < WO

115 120 Val Trp Val Asp Glu Phe Glu Asp Met Asp Leu Leu Asp Cys Leu Asp 135 Ala Thr Trp Ser Gly Ala Ile Val Tyr Val Asp Asp Arg Ser Thr Lys 155 Asn Leu Ser Arg Pro Tyr Ala Arg Val Asn Arg Lys Asn Leu Lys Ser Lys Met Met Lys Lys Cys Val Ser Asn Gly Val Arg Phe His Gln Ala Thr Val Val Lys Ala Met His Glu Glu Glu Lys Ser Tyr Leu Ile Cys 200 Ser Asp Gly Val Thr Ile Asp Ala Arg Val Val Leu Asp Ala Thr Gly Phe Ser Arg Cys Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln Val Ala Tyr Gly Ile Leu Ala Glu Val Glu Glu His Pro Phe Asp 245 Val Asp Lys Met Val Phe Met Asp Trp Arg Asp Ser His Leu Asn Gly 265 Lys Ala Glu Leu Asn Glu Arg Asn Ala Lys Ile Pro Thr Phe Leu Tyr 280 Ala Met Pro Phe Ser Ser Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu 295 Val Ala Arg Pro Gly Leu Lys Met Glu Asp Ile Gln Glu Arg Met Val 315 Ala Arg Leu Asn His Leu Gly Ile Arg Ile Lys Ser Ile Glu Glu Asp Glu Arg Cys Val Ile Pro Met Gly Gly Pro Leu Pro Val Ile Pro Gln 345 Arg Val Val Gly Ile Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met Val Ala Arg Thr Leu Ala Ala Ala Pro Ile Val Ala Asn 375 Ser Ile Val Gln Tyr Leu Val Ser Asp Ser Gly Leu Ser Gly Asn Asp Leu Ser Ala Asp Val Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg 410 Gln Arg Glu Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp 425 Leu Glu Gly Thr Arg Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro 440



Val Pro Phe Gly Leu Ser Leu Phe Ser His Ala Ser Asn Thr Cys Lys 465 470 475 480

Leu Glu Ile Met Ala Lys Gly Thr Leu Pro Leu Val Asn Met Ile Asn 485 490 495

Asn Leu Val Gln Asp Arg Asp 500

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/12121

| | FC170399 | 12121 |
|--|---|---|
| A. CLASSIFICATION OF SUBJECT MATTER IPC(6): Please See Extra Sheet. US CL: 435/189, 193, 233, 252.3, 320.1, 325; 536/23.2 According to International Patent Classification (IPC) or to both | h national classification and IPC | |
| B. FIELDS SEARCHED | | |
| Minimum documentation searched (classification system follow | ved by classification symbols) | · · · · · · · · · · · · · · · · · · · |
| U.S. : 435/189, 193, 233, 252.3, 320.1, 325; 536/23.2 | , , | |
| Documentation searched other than minimum documentation to | he extent that such documents are incl | uded in the fields searched |
| Electronic data base consulted during the international search (Please See Extra Sheet. | name of data base and, where practic | able, search terms used) |
| C. DOCUMENTS CONSIDERED TO BE RELEVANT | | |
| Category* Citation of document, with indication, where a | appropriate, of the relevant passages | Relevant to claim No. |
| WO 97/36998 A1 (UNIVERSITY (PARK) 09 October 1997, see entire No:1. | OF MARYLAND COLLECTION OF MARYLAND COLLECTION OF THE PROPERTY | SE 1-8 |
| | | |
| | | |
| | | |
| | | |
| | | |
| · | | |
| Further documents are listed in the continuation of Box | C. See patent family annex | |
| Special categories of cited documents: A document defining the general state of the art which is not considered to be of particular relevance | "T" later document published after the date and not in conflict with the the principle or theory underlying | international filing date or priority application but cited to understand the invention |
| *E* earlier document published on or after the international filing date | "X" document of particular relevance considered novel or cannot be con | the claimed invention cannot be |
| "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) | "Y" document of particular relevance | the claimed invention cannot be |
| "O" document referring to an oral disclosure, use, exhibition or other means | considered to involve an inver | tive step when the document is such documents, such combination |
| *P* document published prior to the international filing date but later than the priority date claimed | *&* document member of the same p | |
| Date of the actual completion of the international search 02 AUGUST 1999 | 15 SEP 1999 | search report |
| Name and mailing address of the ISA.US Commissioner of Patents and Trademarks Box PCT | Authorized officer | |
| Washington, D.C. 20231 | BRADLEY S. MAYHEW | fui |
| Facsimile No. (703) 305-3230 | Telephone No. (703) 368-0196 | , , |

BNSDOCID WO COCCEPT IS A 210 (second sheet)(July 1992)*

International application No. PCT/US99/12121

| A. CLASSIFICATION | OF | SUBJECT | MATTER: |
|-------------------|----|---------|---------|
| IPC (6): | | | |

C12N 1/21, 5/10, 9/10, 15/53, 15/54, 15/61, 15/63; C12P 23/00; C12Q 1/68

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

Dialog and APS

search terms: 1PP, epsilon cyclase, lycopene cyclase, isopentenyl pyrophosphate isomerase and isopentenyl diphosphate isomerase

THIS PAGE BLANK (USPTO)

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

| Defects in the images include but are not limited to the items chec | cked: |
|---|-------|
| ☐ BLACK BORDERS | · |
| ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES | |
| ☐ FADED TEXT OR DRAWING | |
| BLURRED OR ILLEGIBLE TEXT OR DRAWING | |
| ☐ SKEWED/SLANTED IMAGES | |
| ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS | |
| ☐ GRAY SCALE DOCUMENTS | |
| ☐ LINES OR MARKS ON ORIGINAL DOCUMENT | ` . |
| ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY | |
| OTHER: | |

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)